DOCUMENT RESUME

ED 237 497 SP 023 330

AUTHOR Winne, Philip H.; Marx, Ronald W.
TITLE Students' Cognitive Processes While Learning from Teaching. Final Report (Volume One).
INSTITUTION Simon Fraser Univ., Burnaby (British Columbia). Faculty of Education.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Apr 83
GRANT NIE-G-79-0098
NOTE 258p.; For Volume Two, see SP 023 331.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC11 Plus Postage.
DESCRIPTORS Academic Achievement; *Behavioral Objectives; *Cognitive Processes; Elementary Education; Learning Processes; *Learning Strategies; *Mediation Theory; Psychological Patterns; Student Behavior; *Student Reaction; Teacher Behavior; *Teacher Influence; Teaching Methods

ABSTRACT

Research is reported on the cognitive mediational paradigm which postulates that teachers influence students' learning by causing them to think and behave in particular ways during teaching. Four studies are reported. The first describes five teachers and their students and explores, in classroom lessons, the cognitive processes students used in response to teaching and the cognitive processes their teachers intended them to use. The second and third studies employ analogs of classroom teaching in the form of short videotaped lessons. These sought to determine if elementary school students could be trained to perceive and act on common instructional stimuli and whether these operations would facilitate learning. The fourth study constituted an extension of the second and third studies to regular classroom environments. Three major conclusions are offered: (1) Students and teachers operate in ways that reflect the mediating role of students' cognition in classroom learning; (2) Students can be trained to discriminate instructional stimuli and respond with pre-arranged cognitive strategies; and (3) Students' achievement is partly a function of cognitive strategies they activate in response to instructional stimuli perceived during teaching. Methodological issues attendant to the studies are discussed. (Author/JD)

IPRG

SP 2 4 11 Apr 85 23 7

FINAL REPORT

Students' Cognitive Processes While Learning from Teaching

(Volume One)

April, 1983

National Institute of Education Grant
NIE-G-79-0098

Philip H. Winne
Ronald W. Marx

Instructional Psychology Research Group
Faculty of Education
Simon Fraser University
Burnaby, B.C., Canada
V5A 1S6

FINAL REPORT

Students' Cognitive Processes While Learning from Teaching

(Volume One)

National Institute of Education Grant NIE-G-79-0098

Principal Investigators: Philip H. Winne
Ronald W. Marx

Research Assistants:

STUDY I: Joan Belfry, Dawn Howard, Armine Jeknavorian, Beverly Keizer

STUDY II: Dawn Howard, Beverly Keizer, John Walsh

STUDY III: Chuck McCafferty, Lorne Prupas, Sherry Roberts

ACKNOWLEDGMENTS

We would like to thank the participating teachers and students of the Coquitlam School District and the Burnaby School District. Obviously, without their sincere efforts and cooperation this project would not have been possible. Special thanks go to Tom Wheeler and Dr. Blake Ford for their help in locating participants.

Joan Belfry, Dawn Howard, Armine Jeknavorian and Beverly Keizer all contributed to the sections in Chapter 2 that describe the participating teachers and their classrooms. Dawn Howard, Beverly Keizer and John Walsh contributed to the methods sections of Chapters 3 and 4. Chuck McCafferty, Lorne Prupas and Sherry Roberts drafted portions of the method section of Chapter 5. Beverly Keizer created the scoring manual for the essay tests. Eileen Mallory's superior word processing skills made production of the final report possible.

This final report was completed while Phil Winne was a visiting scholar at the Centre for Applied Cognitive Science, Ontario Institute for Studies in Education and Ron Marx was a visiting scholar at the Department of Educational Psychology, University of Arizona. Both of us were partially supported by leave fellowships from the Social Sciences and Humanities Research Council of Canada.

ABSTRACT

Research reported here is based on the cognitive mediational paradigm for research on teaching, which postulates that teachers influence students' learning by causing them to think and behave in particular ways during teaching. These events may lead to changes in outcome variables, such as scores on achievement tests. Hence, the effects of teaching on learning are mediated by students' behavior and thoughts during instruction.

Four studies are reported. The first describes five teachers and their students, and explores in classroom lessons the cognitive responses students report using in response to teaching, and the cognitive processes their teachers intend them to use. The second and third studies employ analogs of classroom teaching in the form of short videotaped lessons. The objectives of these studies are to determine if elementary school students could be trained to perceive and act on common instructional stimuli and whether these operations would facilitate learning. The fourth study constituted an extension of the second and third to regular classroom environments.

Three major conclusions are offered based on the results of these studies. First, students and teachers operate in ways that reflect the mediating role of students' cognition in classroom learning. Second, students can be trained to discriminate instructional stimuli and to respond with pre-arranged cognitive strategies. Third, students' achievement is partly a function of the cognitive strategies they activate in response to instructional stimuli perceived during teaching. Methodological issues attendent to the study of students' cognitive processes in classroom learning are also discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURE

TABLE OF CONTENTS, VOLUME 2

CHAPTER 1

A Cognitive Mediational Paradigm for Research on Teaching

CHAPTER 1

A Cognitive Mediational Paradigm for Research on Teaching

INTRODUCTION

Early History

In 1897, Joseph Rice surprised the world of educational research with his report of an extensive survey that documented the lack of relation between the amount of time students spent in spelling drill and their spelling achievement. His publication was significant in two ways. First, it demonstrated that propositions about teaching could be potent influences on both theory and practice when they were supported by systematic empirical investigation as opposed to merely individualistic raw experience. While rational and emotional arguments were leveled against Rice's finding, there nonetheless was a robustness to his claims that was unmatched in earlier and more numerous personal accounts about instructional effects. Empirical methods became solidly established as means for investigating teaching and for supporting positions about teaching effectiveness.

The second major impact of Rice's article was to raise the question of why a well-respected and widely practiced teaching method did not produce levels of achievement that were believed to be the inexorable product of such teaching. That an expected effect was not observed in data opened up further issues regarding the epistemological status of then contemporary conceptions of teaching and methodologies for promoting knowledge about teaching.

In the six decades that followed Rice's seminal paper, empirical research that explored relations among aspects of teaching and learning was fundamentally utilitarian. Though many researchers asked why certain instructional practices might influence students' learning, the research was fundamentally a trial-and-error search for any socially permissable teaching practice that fostered students' achievement. Though a causal

theory linking teaching to students' achievement was what researchers sought, the dominant empirical method was correlational. This search for teacher effectiveness, irrespective of why such effectiveness should be observed, held "almost complete dominion over the conceptions that most researchers ... brought to the field" (Gage, 1963, p. 114).

Reviews of the many investigations in this tradition were less than positive. The sought-after causal relations between teachers' acts and student achievement generally were unreliable, weak, and perhaps most disturbingly, sometimes contrary to overpowering commitments about what teaching should be like (see Dunkin & Biddle, 1974). For example, Anderson's (1959) meticulous analysis of the studies about teachers' styles of leadership as potential determinants of students' attitudes and achievement eroded scientific support for a highly-valued position many educationalists held about the superiority of democratic teaching styles. A tenable conclusion in the late 1950s was that research on teaching had been generally uninformative about teaching effectiveness, and that the basis of knowledge it claimed could promise little for the future.

Not all researchers, however, allowed this growing despair about the limited results and seeming valuelessness of research on teaching to incapacitate their search for understanding that might lead to better teaching practice. Two major developments appeared in the 1950s and early 1960s that would sustain both researchers and consumers of research on teaching. The first of these was to dominate the next 15 years of the field's growth. The second, while often talked of and approached, would prove much more elusive.

In sharp reaction to the widespread characterization of instructional events as molar or global, as was evident in descriptions of leadership styles or discovery strategies for teaching, researchers of the 1960s and early 1970s turned toward precise behavioral specifications of what teachers and students did in classrooms. Flanders' (1960, 1963) and Bellack, Hyman, Smith, and Kliebard's (1966) systems for categorizing behavioral

interactions in teaching were representative of an emerging movement toward specificity in the way researchers cataloged variables that might index teaching effectiveness. By 1967, Simon and Boyer were able to collect nearly 70 different instruments for observing classroom interaction. Shortly thereafter, Rosenshine and Furst (1973) conservatively estimated the number of such systems to be in the hundreds. The directions suggested by behavioral observation still dominate much current research (e.g., Gage, 1978; Ryan et al., 1982; Winne & Marx, 1977).

Complementing this more fine-grained view of teaching events was a surge in questions raised about why and how discrete teacher behaviors could affect students' achievement. Two events in particular seem to have been major contributors to the emergence and eventual prominence of calls for theories of teaching. The first was a gathering of educational and psychological researchers that led ultimately to the Taxonomy of Educational Objectives (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956). This taxonomy provided researchers with a greater degree of specificity than was available previously for describing students' learning. This system complemented the growing precision of systems for observing classroom interaction. But this was not the most prominent effect of the Taxonomy.

The psychological language used in the Taxonomy to describe categories of learning practically begged researchers to insert into their models of teaching descriptions about how students would respond cognitively to instructional events. For instance, teachers' questions were classified into hierarchies based on claims about how students responded cognitively to teachers' questions (e.g., Aschner, Gallagher, Perry, Afsar, Jenne, & Faar, 1965). The use of language from the taxonomy clearly invited researchers to speculate about how students' cognitive responses to specific teacher behaviors during teaching might promote or inhibit acts of learning and the products of these acts. Since these cognitive mediators of instructional events were theoretical constructs, their inclusion in models of teaching effectiveness brought concerns for theory into direct contact with commitment-

laden research efforts (Dunkin & Biddle, 1974). Researchers now had a language for describing teaching that was paralleled at least partly by theories of learning. Though distant from classroom teaching in many ways, this infusion of psychological jargon into research on teaching provided a foundation for generating theoretical explanations of teaching effectiveness based on behavioral analyses of teachers' and students' verbal activities during teaching.

The availability of theoretical concepts that the Taxonomy provided may not have been, by itself, a sufficient catalyst for researchers to create and test theories of teaching in research. What seems to have insured that researchers would attempt to use theory in their research were compelling arguments made by Gage (1963, 1964) in the Handbook of Research on Teaching and in the widely circulated National Society for the Study of Education yearbook titled Theories of Learning and Instruction. In proposing a logical division between theories of learning and nascent theories of teaching, Gage gave stature and credibility to research on teaching that availed itself of theory.

In a context of growing support for research in the form of grants and contracts, and rising calls for better education emanating from a rekindled social conscience, theories of teaching were seen as a social necessity and a logical next step for many of the various theories of learning developed by psychologists (see Hilgard, 1956). As Gage (1964, p. 271) saw it:

> Much of our knowledge about learning can be put into practice only by teachers. And the ways which these teachers would put this knowledge into effect constitute part of the subject of theories of teaching. Our position is that practical applications have not been gleaned from theories of learning because theories of teaching have not been developed .... Theories of teaching and the empirical study of teaching may enable us to make better use of our knowledge about learning.

Having shifted from molar to molecular analyses of behavioral events in teaching, and having adopted a strong psychological orientation as a basis for explaining effects of teaching, the

next decade and a half would see researchers focus on generating theory in research on teaching.

The 1970s

Precise specifications of teacher-student verbal interactions on the one hand, and attempts to generate and test psychologically-referenced theoretical propositions about cause-effect relations in teaching on the other hand, received much attention in the research of the 1970s. The former reached an apex of development in the performance-based teacher education movement (see Gage & Winne, 1975). The latter remained, for the most part, rhetorical.

In their extensive review of classroom-based research on teaching that covered roughly the period of the late 1960s and very early 1970s, Dunkin and Biddle (1974, p. 425) characterized the field's greatest problem as a "lack of adequate theories of teaching that would integrate and explain its major findings." In reports of studies, most researchers indeed offered nascent theoretical propositions about teaching effectiveness. Rarely, however, were studies designed in such a way as to test one or another psychologically oriented theory. As a result, only infrequently were data generated from research on which to base strong theoretical propositions. The end of the 1970s was marked significantly by the absence of the theories called for earlier by Gage, and more recently by Dunkin and Biddle! Although the edited volume Research on Teaching (Peterson & Walberg, 1979) offered several chapters that suggested routes to follow in developing theoretical structures, the contributors to this volume did not yet have the resources to achieve the goal set for the field nearly 20 years earlier.

Research on teaching, however, has not stagnated. The major developments in research on teaching during the decade of the seventies were methodological and metatheoretical. In a chapter entitled "Theory Construction for Research on Teaching", Snow (1973) tackled directly the task of developing theories for the field. In a second paper (Snow, 1974), he made several sage points for improving general methodology in experimental research

on teaching. One of the major contributions Snow made in his 1974 paper extended and elaborated a theme he dealt with in an earlier paper (Snow, 1968), namely, a concern for improving the external validity of experiments on teaching by using "representative designs." In his words (Snow, 1974, p. 265-66), "the biggest threat to external validity may come when the experiment does not fit the nature of the behavior being studied and, furthermore, does not include the means for discovering this fact." Since teaching is an activity that takes place in an informationally rich, ambiguous and redundant environment, experiments must be representative of this context if they are to inform theory that can be not only descriptive, but prescriptive as well. Snow's efforts were followed by more substantively specific work by Berliner (1976), Winne and Marx (1977), and Doyle (1978).

Beyond criticisms of prior research, these papers provided concepts and principles that could serve as stepping stones for conceiving, conducting, analyzing, and interpreting research on teaching. In his summary of "Impediments to the Study of Teacher Effectiveness," Berliner (1976) highlighted a number of concerns he had for improving the quality of research within the process-product paradigm. Among general concerns for statistical analysis, the nature of instruments that could document the status of independent and dependent variables, and methodological problems regarding issues such as internal validity, he included a section on the mediation of teacher effectiveness through student behavior. Two basic messages appeared in this section. The first was that "We are now convinced that the mediating link so necessary to consider is a student's active time-on-task" (Berliner, 1976, p. 10) and that "some variables thought to be quite important by educational theorists are in fact unimportant to, unperceived, or unperceivable by students" (Berliner, 1976, p. 11). In our view, the latter question is a theoretically richer formulation of the former. Gage (1978, p. 76) put it this way: "Academic learning time, in the form of allocated and engaged time, is, in a sense, a psychologically empty quantitative concept. We need better analyses of how that time is filled, of

what learning processes go on during academic learning time."

Implications for the Project

This brief history of research on teaching highlights several key aspects regarding the current status of theories of, and research on, teaching. First, there is now full agreement that systematic empirical methods are essential if research is to provide data from which theories can be created, tested, and revised.

Also, there is widespread, but not complete, agreement that relatively precise analyses of the behavioral environment of teaching are appropriate for this task (cf. Bronfenbrenner, 1976). Third, theories of learning provide a promising point of departure for the development of theories of teaching, though we believe that these are but one family of theories that will prove useful (e.g., see also Gall & Gall, 1976). Finally, there is a recent but nonetheless forceful call for theory-oriented research to be conducted in a manner that represents rather than distorts the phenomena of interest (Snow, 1973).

Not all intellectual developments achieved in this history were without cost, however. In particular, Gage's call for links between theories of learning and theories of teaching have not yielded well-articulated, empirical ties between the two varieties of theories. Rarely has research on teaching been used as a basis for generating or altering propositions contained within theories of learning. And, beyond a fast-developing theory of learning outcomes and procedures for testing them (see Gagné & Beard, 1978), little empirical verification can be found for descriptions of how instructional events bring about student learning in terms of theories of learning (Winne & Marx, 1977; Winne, 1982b).

The cost to theories of teaching wrought through previous attempts to join them with theories of learning has been one of losing sight of fully one-half of the unit that must be considered in the teaching-learning process. Specifically, sets of "teacher should" principles (Gage, 1980) comprising contemporary prescriptive theories of teaching speak only of the teacher. They ignore the psychological context within which the effects of

teaching occur, namely, the everpresent thinking in which students engage as they are attempting to make sense of and respond to the instructional activities orchestrated by the teacher (Winne & Marx, 1979). At present, the box that frames students' cognitive activities during teaching is relatively empty (cf. Winne, 1983c).

Recently the process-product paradigm for research on teaching as described here has undergone a logical revision. Notable among updates to this paradigm are what Doyle (1980) calls the "mediating process paradigm," or what Winne (1982) refers to as the "cognitive mediational paradigm." Both views postulate that teachers do not directly influence student product variables, such as achievement. Rather, teachers influence students by causing them to think and behave in particular ways during the course of teaching. These intra-instructional events, in turn, may lead to changes in outcome variables. Hence, the effects of teaching on learning are mediated by students' behaviors and cognitive processing during instruction.

The cognitive mediational paradigm opens up to question the match between researchers' hypotheses about how learners react mentally to teaching events and the ways learners actually process information cognitively in classrooms. For example, consider a teacher asking an application question during instruction (Hunkins, 1972). Although such teaching events can be observed reliably in terms of their surface behavioral characteristics, characterizing students' cognitive responses to them raises crucial issues. If students fail to demonstrate mastery of the content called for by an application question on an achievement test, or if they do not use the process of application with new curricular content, researchers usually concluded that the teacher's application questions did not promote learning. Researchers' reasoning about this chain of events, however, has not included an empirical test of whether, during teaching, students ever engaged in the psychological processes that researchers assume to accompany application questions. Without such data to document students' mediations of the teacher's questions or other supposedly instructional events, descriptions

about how instructional variables affect students' psychological processing of information, and ultimately their achievement, are fundamentally speculative.

Failing to document that students use the particular cognitive responses researchers associate with a given instructional variable allows several rival hypotheses to be proposed when treatments fail to affect students' learning positively (Winne, 1979, 1982a). Consider again the case of a teacher asking an application question where some students do not demonstrate achievement on a test item that is directly related to the answer of the question asked in class. First, although the teacher's application question was apparent to a researcher, some students may not have been attending to the teacher at that point. In effect, these students do not experience the instructional event. Another possibility is that students may have heard the question, but may not have understood how to respond cognitively to it beause the cognitive processes the teacher intended were not communicated clearly. Other students may have heard the question and understood the kind of cognition intended by it, but may have lacked either the required knowledge to insert into cognitive processes or the ability to draw on cognitive processes perceived to be the kind the teacher intended. For these students, the application question could not function as the teacher, or the researcher, intended. A fourth possibility is that some students noticed, understood, and were able to carry out the cognitive work to answer the teacher's application question but, for some reason, chose not to do it. These students lacked motivation or thought they knew a better way to reach an answer to the question. If each of these four possibilities characterized only two or three students in a class, at least one-third of the class would not be cognitively responding to the application question in a way theorized by the researcher or intended by the teacher. A recent study by Dillon (1982) empirically confirms that, at least for each single student who does provide an overt, observable answer to a teacher's question, the odds are 50-50 that the answer reflects cognition

attributed to the question by the researcher. Thus, it would not be surprising that some students failed to learn what was expected, or that mean achievement in such a class was not higher than a comparison class which was not asked application questions (Winne, 1979).

We postulate that some teacher behaviors directly influence students' cognitive processing of content during teaching. These cognitive processes, in turn, determine whether students learn in accord with teachers' intentions and researchers' models of how learning takes place. Since these cognitive responses are central in determining "what is processed, how it is processed, and therefore what is remembered" (Rothkopf, 1976, p. 116), they are essential elements in theories of teaching and in research derived from these theories.

If the cognitive mediational model is correct, research on how teaching affects learning will need to change. Specifically, in many previous investigations researchers have stressed teacher behaviors and, to a slightly lesser extent, student behaviors. Both were treated as direct causes of students' learning. The cognitive mediational model will require that research explore how students perceive learning tasks (Weinstein, in press; Winne & Marx, 1979) and what they do cognitively to bring about learning based on these perceptions. Such research also would need to link teachers' decisions about how they intend to direct the cognitive processing students should use to learn from teaching (Shavelson, 1981), teacher behaviors that communicate these intentions to students (Peterson, Swing, Braverman & Buss, 1982), students' perceptions and use of cognitive processes to learn from teaching (Marx, Winne, & Howard, 1982), and students' behaviors that inform teachers how well students are progressing (Marx, 1978).

Rich theories and their attendant bodies of empirical findings exist to help fill these lacunae. They exist partly in the literature on the psychology of learning. Along with Gage and others (e.g., Calfee, 1981; Doyle, 1978, 1980) we think it necessary to draw upon this literature, specifically the burgeoning research and theories of cognitive learning, to inform

and enrich theories of teaching. To do this so as to advance prescriptive instructional theory, we also believe that research questions must be cast and methodologies must be developed in the spirit of Snow's (1974) call for representative design. But, before this research is begun, it should be documented that the cognitive mediational paradigm is worth pursuing. We turn to this task in the next section.

VIABILITY OF THE COGNITIVE MEDIATIONAL PARADIGM

The models that guided research on teaching over the last few decades have undergone significant transition. In the last section we provided an overview of changes regarding the adoption of empirical methods, the sharpening of focus for viewing teaching events, the call for links to theories of learning, the argument for representative methodology, and attempts to integrate these to build foundations for generating theories about teaching. In this section, we present a more detailed picture of the transition from a process-product paradigm that characterized the vast majority of previous research on teaching to a cognitive mediational paradigm. This newer paradigm elaborates and focuses more precisely on questions about relationships among instructional events in classrooms and students' achievement by making use of theories of learning.

Research on teaching conducted in the last two decades was strongly influenced by conceptions put forth in the Handbook of Research on Teaching (Gage, 1963). Doyle (1978) recently summarized the process-product paradigm that characterized a synthesis of those views. Essentially, this paradigm was concerned with predicting the degree to which particular teaching events would correlate with, and ultimately control, the products of teaching defined as students' attitudes and achievement. But, as we demonstrated in the previous section, this focus on teacher behaviors delivered during instruction and measures of student learning obtained after instruction has begged the question about the psychological responses students made to teaching during instruction.

The influence of methodological behaviorism (e.g., Skinner, 1953) helped remed'ate this deficiency by encouraging researchers to document more precisely the behaviors in which students engaged during lessons. Thus, some observation instruments employed in process-product research on teaching began to catalog both teacher and student behaviors in the context of teaching. But, very little data has accrued that has validly linked students' behaviors to the cognitive processes that researchers favored as the explanatory mechanisms to account for teaching effects.

As process-product research began to yield findings that violated researchers' commitments to speculative psychological mechanisms linking behavioral events in the classroom to tudents' subsequent achievement, questions were raised about whether the process-product paradigm could illuminate the means by which students' achievement might be influenced during teaching. Doyle (1978) and Winne and Marx (1977) summarized these questions in their reviews of research up to that time. Doyle focused fundamentally on the three paradigms labelled process-product, mediating process, and ecological. Winne and Marx argued for including in research on teaching perspectives that took into account teachers' cognitive processing, students' cognitive processing, and bidirectional behavioral interactions that were the medium for teachers and students to communicate about cognitive processing students could use to learn from teaching.

Doyle's and Winne and Marx's characterizations of deficiencies in prior research on teaching raised questions about how validly behavioral events occurring in classrooms were mapped onto theoretical constructs which these behavioral events were believed to entail. For example, in his discussion of the mediation of teacher effects through student behavior, Berliner (1976, p. 10) stated that "intermediate links in the causal flow (between teacher behavior, students' cognitive processing, and subsequent learning) require us to examine the students' attending and information-processing behavior." In other words, the process-product paradigm was not representative of the range of phenomena that characterize teaching and learning. "Information

about students (was) confined primarily to scores on pre-test and post-test achievement measures. Even when data about students classroom behaviors are available, the tendency in the past has been to deemphasize this evidence in interpreting findings" (Doyle, 1978, p. 167).

The Problem of Students' Cognitions

Having raised the question about the role of students' cognitive processing in learning from teaching, there still remained considerable work to be done to focus this question and to generate methodologies for investigating phenomena related to it. In approaching these two objectives, considerable light was shed on the difficulties which plagued research on teaching conducted under the guidance of the process-product paradigm.

The fact that researchers had to ask the question of how students respond cognitively to teaching events suggests that they had been overly optimistic in assuming an isomorphism between the ways in which they thought learners were reacting mentally to teaching events and the ways in which learners were actually thinking in classrooms.

It is not surprising that students' achievement failed to develop as expected if they had not engaged in the cognitive processes researchers were attempting to promote by teachers' use of different teaching behaviors. These same cognitive processes, however, constituted the foundation on which researchers were attempting to create theories of teaching. When a researcher's assumptions about the occurrence of these processes does not correspond to their actual use by students, theories of teaching necessarily will suffer from a lack of construct validity of treatments (Cook & Campbell, 1979). This was what plagued research that grew out of the process-product paradigm. But, a logical argument about the mediational role of students' cognitive processes in learning from teaching is not sufficient for adopting the cognitive mediational paradigm. Empirical questions remain as to whether these phenomena exist and whether their importance can be demonstrated.

Evidence for the Cognitive Mediational Paradigm

In order for the foregoing argument about the mediational role of students' cognitive processes to be supported empirically, it is necessary to demonstrate at least four things. First, there is no doubt that students can learn from teaching as presently delivered in classrooms. In addition, a fundamental assumption of cognitive psychology is that learners actively construct mental representations of their environment, rather than passively react it. Presumably, in an environment such as a classroom where the activities are purposive, this active construction of knowledge and skills is bounded by the goals of instruction. That is, one would not expect learners to construct cognitive representations of the content of instruction that are not in some way related to the teacher's presentation of this information. Thus, a first necessary condition for accepting the validity of the mediating process paradigm is that students perceive the occurrence of specific instructional events. If such perceptions are not attained, then it would be unlikely that students would engage in the deliberate use of cognitive processes that the teacher intended in response to these events.

Many areas of research about teaching provide evidence that students notice discrete teacher behaviors. In a highly controlled experiment on the effects of teachers' use of structuring, soliciting, and reacting behaviors, Winne (1977; see also, Clark, Gage, Marx, Peterson, Stayrook, & Winne, 1979) used aptitude-treatment interaction methodology to examine aptitudes related to whether students attended to their teacher's use of these behaviors. There was clear evidence that students' attention to teacher behaviors that made up the experimental treatment were associated with the actual occurrence of those events and with students' aptitudes. For example, with respect to a cluster of inter-related teacher behaviors constituting the dimension of reacting, students' attitude toward the subject matter, their general ability, and the actual degree to which the teacher engaged in reacting behaviors predicted the extent to which students characterized their teacher as having used these

teacher behaviors. Thus, Winne found that, in addition to the actual teacher behaviors used during instruction, students' perception of the occurrence of those instructional events during teaching was associated with their aptitudes.

In another context, Weinstein, Middlestadt, Brattesani, and Marshall (1980) documented that students perceived their teacher to use several specific teacher behavior variables differentially when interacting with high and low achieving students. For example, students indicated that, relative to high achievers, low achievers tended to be the recipients of negative feedback and to be directed more by the teacher while performing their classroom activities. Also, high achievers were judged by students to be given more choices while accomplishing their academic tasks, and to receive statements by the teacher indicating that more was expected of them in academic performance.

In contrast to Winne's (1977) study, this one suggested that students also can perceive the existence of relatively global teacher behaviors in the context of teaching. These two studies jointly illustrate that students do attend to occurrences of particular teacher behaviors during instruction. The fact that Winne found students could recognize the occurrence of specific behaviors embedded in a large facet of teaching suggests that students may be able to perceive both discrete events as well as clusters of related activities in teaching. That students' aptitudes influenced the degree to which they noticed these events suggests that conditions under which students may notice these events could vary. The study by Weinstein et al. documented that students also can perceive global characteristics of classroom teaching. Thus, the level or complexity of teacher behaviors investigated by the mediating process paradigm probably can span the range of instructional events currently being researched in the field of teaching effects.

A second necessary question to be asked of the mediating process paradigm is whether students' perception of the occurrence of teacher behaviors actually relates to their subsequent achievement. A study by Stayrook, Corno, and Winne (1978)

provides evidence on this point. Using part of the data from Winne (1977) and Clark et al. (1979), path-analysis techniques were used to analyse the links between students' aptitudes, manipulated occurrences of specific teacher behaviors, students' perceptions of the occurrence of those teacher behaviors, and subsequent student achievement. Stayrook et al. found that, beyond the effects that students' aptitude and teachers' actual use of specific teacher behaviors had on subsequent learning, students' perceptions of the occurrence of particular teacher behaviors also were directly linked to their levels of achievement. Stayrook et al. noted that "the mediating effect of student perceptions may be behavior-specific. For structuring and reacting, it seems that such perceptions do act as mediating variables, but this is not the case for soliciting" (Stayrook et al., 1978, p. 55).

This study is important in two respects. First, it provides evidence that students' perceptions of the occurrence of teacher behaviors can be related to achievement measured at a later point in time after teaching is over. Additionally, the study also documented that the link between students' perceptions of particular teacher behaviors and their subsequent achievement may vary for different kinds of teacher behaviors. Descriptions of the nature of this variation constitute an important goal for research guided by the mediating process paradigm.

The third key question with which the mediating process paradigm must wrestle is whether students understand that a teaching event is intended to engage them in one, as opposed to another, particular cognitive process. If students attend to particular teacher behaviors, and if these perceptions are related to achievement, it still remains to be demonstrated that students engage particular cognitive processes in response to teacher behaviors. Studies by Morine-Dershimer and her colleagues (Morine-Dershimer, 1982; Morine-Dershimer & Fagal, 1980) provide some evidence on this point.

These researchers videotaped students participating in short lessons. Using the videotape as a stimulus to cue recall, they

interviewed the students about their participation in the lessons. The students' responses to queries about the functions of teachers' questions were categorized according to what the students understood about why the teacher was asking particular questions. For example, 36% of the teacher's questions about which students were asked in these interviews were classified by the students as providing an opportunity for the teacher to teach further information while students were responding. To speculate on the cognitive processes students might believe were intended in this context, students might believe that questions were intended to create for them a set to be used for receiving information following the question itself.

Students were also asked about their understanding of the functions of praise. In 59% of these incidents, it was reported that the teacher's use of praise indicated that a student had a correct response or a good idea. Again, speculating on cognitive processes that students might use in this context, praise could be a signal either to store in memory the immediately preceding idea, since praise cued its importance, or to identify it in some special way because it would be useful later in the lesson. In either case, from the point of view of the cognitive processing of students, praise can be seen as carrying informational value as much as, if not more than, carrying motivational value (cf. Brophy, 1981).

In addition to other findings in these studies, there exists evidence that students infer meaning for their cognitive processing as a function of the occurrence of particular kinds of teacher behaviors. However, there still remains a fourth question which must be asked in the context of the mediating process paradigm in order to accept it as an appropriate characterization of students' learning from teaching. If students attend to the existence of teacher behaviors, if their perception of these behaviors is related to achievement, and furthermore if students interpret these behaviors as signals that certain cognitive processes be engaged, one still must be able to demonstrate that when all three of these events occur, learning can be influenced

when students' cognitive processing is controlled in response to teacher behaviors. The issue here is whether students' cognitive processing of content, when cued by teacher behaviors, can become an independent variable in research on teaching. This is essential in order to propose and test theoretical claims about causal links between students' mediating cognitive responses to teacher behaviors and students' achievement.

Three studies can be cited that provide preliminary evidence that the answer to this question is affirmative. Koopman and Newtson (1981) performed an experiment in which they instructed university students viewing a videotaped lesson to focus either on the smallest steps of the lesson that seemed natural and meaningful to them, or the largest steps of the lesson that seemed natural and meaningful. These researchers attempted to control participants' perception of the videotaped teacher's behavior by instructing them to focus at either an atomic or a molar level. Koopman and Newtson found that these simple instructions affected the participants' description of the instruction. Thus, students watching a videotape as if they were participants in a lesson were able to control the level at which they attended to teacher behaviors. Also, there was an indication that lessons which encouraged relatively atomic perceptual analysis produced greater learning than lessons that encouraged more molar analysis. Overall, then, these findings suggest that, not only is it possible for students to control their attention to teacher behaviors, but that certain kinds of cognitive responses to teaching are associated with achievement. However, since this was an analogue study rather than a direct test of an instructional effect, it failed to document a clear relationship between manipulating students' cognitive processing in response to teacher behaviors and subsequent learning.

Two other studies demonstrated that it is possible to influence students' cognitive processing in response to instructional cues, and that such influence affects subsequent achievement. In a study using university students in a lecture situation, Winne and Marx (1980) successfully instructed students

to make specific cognitive responses to the occurrence of specific lecturer behaviors (e.g., giving a new example of a concept). Contrary to expectations, attempting to change these relatively sophisticated learners' cognitive responses to events during a lecture interfered with learning rather than facilitated it. Though the observed effects of training on students' achievement were in a direction opposite to the hypothesis, this study nonetheless provides evidence that it is possible to control students' cognitive responses to teacher behaviors, and that such control does influence learning.

A later study by Winne (1982b) attempted to correct several of the flaws in the preceding experiment. By providing students greater practice in the cognitive responses to particular instructional stimuli, he reasoned that students would be able to execute them relatively expertly. Using upper elementary students who read a text augmented with either instructional objectives or adjunct questions, Winne found that it was relatively easy to instruct learners to make particular cognitive responses to these text-based instructional stimuli. Using aptitude-treatment interaction analyses, he also found that training students to use cognitive responses upon encountering instructional objectives before reading a text interacted with students' general ability. Achievement was enhanced for low ability students who received direct instruction about the cognitive responses to make in response to instructional objectives. The general conclusion drawn by Winne was that students could be trained to make specific responses to instructional stimuli, and that, with respect to subsequent achievement their prior aptitudes sometimes influenced the effects of such control over their cognitive responses.

Conclusion

The foregoing brief review of illustrative studies shows that all four empirical questions that must be asked of the cognitive mediational paradigm can be answered affirmatively. Thus, there is empirical evidence to support the proposition that research in the service of theories of teaching can be linked to theories of learning by focusing on the cognitive responses that students make

to teacher behaviors. Given the logical argument and the available empirical evidence justifying the viability of the cognitive mediational paradigm, we judged that an exploration of teaching as mediated by students' cognitions was appropriate.

The series of studies undertaken in this project focused on several issues. First, what intentions do teachers hold for their students' thinking during instruction? What is the range of these cognitive processes and how do teachers working in their classrooms communicate these intentions to their students? Second, what teacher behaviors do students perceive as conveying information about how to think during teaching, and how do they interpret these signals? Third, can students be taught to perceive pre-arranged signals in teachers' behavior that communicate intentions for student thinking, and to carry out particular cognitive learning strategies as responses to these signals in the context of classroom teaching?

We approached these three questions in a series of studies that began by studying natural teaching in intact elementary school classrooms. Our objectives for this first study were to document teachers' views and students' views about teaching as a means for influencing students' cognitive processing of curricular content. In particular, we asked teachers to describe the intentions they had for students' cognitive processing at varying points in lessons, and to identify the behaviors they used to communicate these intentions to their students. Then, to cross-validate and extend the picture of students' cognitive mediation of teaching as painted by teachers, we interviewed students to discover whether they were aware of the behavior that their teachers had described, and whether their understanding about how curriculum was to be processed cognitively matched the teachers' intentions. In addition, we sought to identify conditions of teaching and of students' cognition that influenced the degree to which students believed they could carry out the cognitive processing they perceived their teacher intended.

The next phase of the project experimented with procedures that would modify students' perception of instruction and increase

their abilities to execute cognitive strategies we hypothesized would enhance their achievement. Two controlled experimental studies were conducted in which students participated in lessons by viewing videotapes of highly structured recitations. Although most students had worksheets available during these lessons that summarized the material to be learned, only some of them had been trained previously to coordinate instructional cues delivered during the teaching with cognitive processing and written manipulations of information to be tested. Thus, these studies explored the degree to which students' intra-lesson cognitive and overt responses to teaching could be directed by signals given during teaching, and whether these mediations of teacher behavior were related to achievement.

Building on the methods and findings of the three preceding studies, a series of five concurrent parallel experiments were carried out in the third phase of this project. These experiments involved careful observations of participating teachers' natural use of cues for students' cognitive responses. Based on these observations, a program was devised to train some students in each of the five elementary classrooms to respond to their own teacher's signal for carrying out an adaptable sequence of cognitive and written activities. There were three purposes for these studies. The first was to test whether procedures that characterized the earlier highly controlled experiments could be adapted to the dynamic milieu of recitation lessons delivered under the conditions of everyday teaching in elementary classrooms. The second purpose of these studies was to explore for influences that students' aptitudes might have on their learning. In particular, we focused on contrasting students who used their naturally developed responses to instruction with others who were trained to respond to teaching as their teacher intended. The third purpose of these studies was to investigate whether students who had been trained to make certain pre-arranged cognitive and written responses to their teacher's signals learned more than their untrained peers using untampered-with learning strategies.

CHAPTER 2

Students' and Teachers' Thoughts about Strategies for Learning from Teaching: Study I

CHAPTER 2

Students' and Teachers' Thoughts About Strategies for Learning from Teaching[1]

INTRODUCTION

One of the major objectives of this project was to describe aspects of the cognitive mediational paradigm under conditions that represented rather than distorted the behavioral and cognitive ecology of classrooms. Thus, the source of data for this study was teaching in schools as it was created by the teachers and students participating in that teaching, and these same participants' subsequent analyses of their involvement in that teaching.

The general procedure followed to obtain these data was to videotape teachers' lessons. Following as close as possible upon a lesson, ususally within 10 minutes, the teacher was interviewed by a research assistant who had been trained to elicit the teacher's analysis of instructional events in terms of the cognitive mediational paradigm. To insure as much as possible that these analyses were reflecting actual events in the preceding lesson, the primary stimuli for cuing the teacher's analyses of these events was the videotape of that lesson.

A small group of students was interviewed as immediately as possible after the teacher interview about a representative sample of incidents that the teacher had analyzed because he saw them as instructionally important. Incidents were chosen to obtain some from each of the beginning, middle, and end of the lesson. Again, to anchor the students' analyses in the actual lesson that they had just experienced, the videotape served as the primary cue about what occurred during teaching. Audiotapes of these two interviews became the raw data for our analyses of the cognitive mediational paradigm as it was represented in the lessons sampled for this project.

[1] An article based on this Chapter appeared in the Elementary School Journal, (Winne & Marx, 1982).

The purpose of this study was to explore the cognitive responses students made to teaching, and the cognitive processes their teachers intended them to use in typical classroom lessons. One objective was to document how students and teachers saw cognitive processes as mediators between teaching events and learning. When such mediations were reported, our second objective was to identify whether students' cognitive activities were congruent with their teacher's intentions for how students should think.

Documenting these mediational phenomena in their natural setting provides a possible link between theories of cognitive learning and theories of teaching. In this regard, a table of the relative importance or frequency of all the cognitive processes intended by teachers and engaged by students is less important than evidence that such events occur. Thus, frequency counts and distributions of occurrence of these phenomena are neither central to our purpose nor appropriate. As a consequence, this study departs substantially from typical classroom observation studies that have had as their goal the quantification of variables and the determination of statistical relationships among them. Moreover, because we cannot be certain that all possible categories of instructional events were represented in our sample of lessons, the data reported in this first study should not be viewed as an exhaustive list of the cognitive mediational links between classroom teaching and learning. At this point in research about the cognitive mediational model, if we can demonstrate the occurrence of cognitive mediational links between teaching events and learning, we believe that experimentation based on the paradigm is warranted. We leave to future research the task of documenting a representative "state of nature" regarding students' cognitive processes during classroom learning, and the means by which teachers signal their intentions about these cognitive processes to students. However, in the studies to be reported later, we will take up questions about the actual effects of students' cognitive mediations of teaching on their achievement.

METHOD

Participants

Teachers. Three male and two female teachers from five different schools in two suburban school districts in the metropolitan area of Vancouver, British Columbia volunteered to participate in the study. Two taught grade seven classes, two taught grade five classes, and one taught a grade four class. Upper elementary grades were selected for two reasons. First, children at these grade levels generally have developed sufficient verbal and intellectual skills so that self-report procedures could be used productively. Second, at these grade levels, most curriculum areas are taught by the same teacher, thus providing a basis for generalizability across teachers, students, and curriculum topics.

For reasons of confidentiality, fictitious names are used in this report for all the teachers. To further guarantee confidentiality, all references to the teachers will involve male pronouns (determined by a coin toss).

The following descriptions of the teachers and their classrooms are intended to paint the rich context within which the data to be described later were collected.

Mr. Acton. Students in his class normally were seated in groups of four at tables organized in V-shaped rows. Some students had a clear view of the front of the room and the blackboard, while others had to turn in their seat whenever that part of the room was used for instruction. Students had their own places, but they were free to move about the room as the lesson required. There were also three or four areas around the side and back of the room for small group activities such as art, reading, and music.

Only twelve students were involved for most lessons. At these times, a student teacher would take the rest of the class to anot er room for instruction. Two of the lessons we videotaped involved the entire class after the student teacher had left the school.

Being in the school is like being in a beehive. There seems to be continuous activity. Teachers and children move about the building throughout the day producing a fairly high ambient noise level. As head teacher, Mr. Acton has some administrative responsibilities and was acting principal when the regular principal was absent. He is an outgoing person who is involved in many school activities.

During instructional sessions, Mr. Acton maintained a relaxed atmosphere within fairly highly structured activities. Students were well aware of what was considered appropriate classroom behavior. Basic objectives and methods for lessons were planned, but classes remained flexible so that, if they desired, students could explore topics in more depth. Mr. Acton demonstrated an awareness of the students' learning abilities and styles, and encouraged all students to participate in the lessons and help one another with learning tasks.

Mr. Anderson. Mr. Anderson has taught both elementary and high school. He is very involved in school projects, and assumed some responsibilities for the principal. He also is involved with his students in several extra-curricular activities such as sports and dancing.

Only a portion of the class participated in a lesson while the others went to the library with a student teacher. The students' desks were in rows facing the blackboard, and the class was decorated with students' work from various projects.

Mr. Anderson planned his lessons in detail and referred to his written plan frequently during teaching. His Masters degree work in using audiovisual equipment in instruction was apparent in his lessons. His lessons began as teacher centered in that he set the goals, determined the activities, and did most of the talking. As the lesson progressed, he used questioning techniques to involve students in discussing the topic and to check their understanding. This sometimes evolved into student-centered activities. Mr. Anderson's lesson formats varied predictably according to the subject being taught. For example, science involved students in experimenting, while health was a lecture-

note taking format. He usually checked with students individually as they worked, giving them corrective feedback.

The atmosphere of the class is characterized by a busy mood with students actively participating in lessons. Mr. Anderson is friendly and encouraging to students, often joking with them. He knew them well, and used his knowledge of their abilities and interests in determining who to call on during lessons. Students are expected to and do assume responsibility for their behavior so that the class is generally cooperative and well-disciplined.

Mr. Lehmann. Mr. Lehmann is a personable, authoritative and outspoken individual who takes pride in his work and seems committed to maintaining a high standard of performance in his teaching skills and strategies. He had taught for many years at the high school level before making the transition into elementary school, where he now teaches a grade five class of eighteen students.

The school operates a French immersion program in which students are educated in the two official languages of Canada from kindergarten to grade seven. All of the students in Mr. Lehmann's class, with the exception of one child who was a new immigrant, were fluent in both French and English. The classroom is in a "temporary" structure situated about 12 meters from the main building. There are windows facing the playground on one side of the room and blackboards on the opposite side. The student desks are arranged in rows in the middle of the room facing the blackboards. The room is filled with student drawings posted on the walls and their poems hang from the ceiling. In one corner of the room, videotaping equipment is set up for use by Mr. Lehmann and his students.

Mr. Lehmann maintains a very structured and orderly classroom. The climate can be characterized as one of mutual respect, with both the teacher and the students enjoying their days together. He plans and prepares lessons ahead of time and adheres quite closely to an instructional process he acquired

while taking in-service courses on Instructional Theory Into Practice. The sequence of activities used in all his lessons included the following: sponge exercise, anticipatory set, modeling, guided practice, monitoring, time framework, raising level of concern, active participation by students, paraphrasing by students, accountability, and closure.

During the first two weeks of school, Mr. Lehmann taught these procedures to his students and explained their purpose. As a consequence, the students are able to follow Mr. Lehmann's instruction and understand his expectations regarding their behavior in the classroom.

When teaching a lesson, Mr. Lehmann assumes responsibility for explaining, directing and goal-setting. Halfway through the lesson, these responsibilities shift to the students who are held accountable for their learning and who are expected to participate actively in the lesson, giving and learning through paraphrasing.

Mr. Orr. The 29 students in Mr. Orr's grade five class sat at individual desks placed around the perimeter of the room. Most students had a clear view of the blackboards situated at the sides of the room. The open center area of the room was used for most of the formal instruction and for small group activities. A large room adjoining the classroom was also available for art and small group activities. Bulletin boards, walls, and doors were covered with the results of students' projects on a variety of topics.

Twenty-five of the students participated in the research project. Most of the videotaped lessons involved the total class. For a few lessons, those not involved were given separate instruction in another room by a student teacher.

Mr. Orr's lessons were planned to combine presentation of material, student clarification of new information, and application in a task carried out by individuals or groups. The teacher displayed a keen awareness of each student's learning abilities and needs, and an obvious enjoyment of teaching and his students. His special interest in art and audiovisual materials meant that lessons were colorful and that a variety of visual aids often were used. Although the lessons were quite structured, Mr. Orr showed flexibility in allowing students to explore concepts or

ideas in which they showed interest, even when this had not been planned as part of the lesson.

During the lessons, he moved about the room encouraging individual students to participate, checking that those who normally had difficulty were grasping ideas, and giving a great deal of positive responses to all of the children. The students appeared happy and participated eagerly in learning activities.

Mr. Richmond. Mr. Richmond is a high energy, verbal individual whose enthusiasm for teaching is apparent. Seeming to have barely enough hours in the day, he has an active professional life outside the classroom with various involvements in sports and teacher inservice.

In the classroom he stresses the need to attract students' attention, almost seeing himself as an entertainer as well as a teacher. He is very aware of student's feelings and abilities, and uses questioning and encouragement techniques frequently. While he is friendly to his students, he also is clearly a disciplinarian. He expects students to act responsibly, often leaving them to work unsupervised. They respond well to Mr. Richmond's trust.

Mr. Richmond's planning for each lesson is detailed. Lessons follow a regular format of relatively teacher-centered recitation shading into lectures. These are followed by discussion or small group activities where Mr. Richmond makes himself available to answer questions and clarify assignments. He also uses this time to monitor the progress of individual students and provide personalized feedback to them. His interactive style is usually democratic, although he assumes total responsibility for deciding standards of class work and choosing curriculum content.

Students. A total of 113 children participated in 149 interviews over the course of the study. Informal consent was obtained from the parents or legal guardians of all participating children. These interviews were conducted by individual research assistants with groups of 2-6 students. Children were randomly selected by their teachers to participate in the interview sessions, and all participating children in each of the five

classes were interviewed at least once. Seventy-six children were interviewed twice and five were interviewed three times during the ten weeks of data collection. Since we attempted to have at least three children in each interview, some children participated more than once due to small class sizes in some cases. A tabulation of the number of students interviewed in the different classes and grade levels is shown in Table 1.

Sources of Data

Videotapes. Fifty classroom lessons were videotaped over the period of the study. Ten tapes were made of lessons by each participating teacher. The lesson content varied across mathematics, science, language arts, and social studies. The length of the lessons ranged from approximately 25-45 minutes. One of the 50 videotapes was lost due to equipment failure.

Audiotapes. Audiotapes were made of the teacher interviews and the student interviews. For the teacher interviews, 48 recordings were made in total. One tape did not record because of machine failure; a second was lost. For the student interviews, 49 audiotapes were successfully completed, one was not recorded due to machine failure.

Interview materials. Logs were kept by the research assistant during the videotaping of each lesson. Part of a log entry was a meter reading on the videotape corresponding to teacher behaviors which the research assistant judged as likely to invite or otherwise involve students in cognitive responses at that point in the lesson. To identify these incidents, the research assistant had to be able to postulate a reasonable or possible mediating cognitive process in which students might engage in response to the instructional event. Preparation for this task was accomplished through simulations prior to the start of data collection. Along with us, the research assistants watched videotapes of lessons used in our teacher education program. Together, we speculated about mediating cognitive processes that the teacher might be signalling, and discussed the reasonableness of these speculations. A brief behavioral description of the event and the research assistant's speculation

TABLE 1

Distribution of Students Interviewed in Each Class

Teacher	Grade	Number of students interviewed	Number of students interviewed more than once
Acton	7	12	7
Anderson	4	27	8
Lehmann	5	17	12
Orr	5	24	3
Richmond	7	33	1

about the teacher's possible intentions for students' cognitive mediational processing completed each entry in the log. These events were candidates for analysis during the teacher interview.

For the student interview, a structured format, reproduced in Figure 1, was designed to standardize the questioning strategy used by the research assistants. The student interview schedule was constructed to probe in detail students' understanding of their cognitive mediational processing during instruction while accomodating their inexperience in describing these sorts of events. At each level the interview schedule was designed to elicit as much specificity as possible about the student's cognitive processing in the student's own words. When students were unable to do this, the interviewer suggested several mediating thought processes from which a student might choose to elaborate (for example, see Level II, Figure 1).

Following an introduction, the schedule began at Level I by soliciting students' perceptions about relatively global cognitive states and behavioral events in the lesson. As the interview proceeded to Level II, the research assistant framed questions that probed more deeply students' perceptions of the cognitive processes possibly supporting their participation in instruction. Following this level, the questions sought information about qualitative features of the instructional events surrounding their cognitive mediational processing (Level III), and students' comprehension of how their cognitive mediational processing might relate to acquiring and retaining the content of the lesson (Level IV). Finally, the students were asked to describe alternatives to the instructional event and its corresponding cognitive mediational processes (Level V); and to express their view about how the instructional event was integrated within the flow of the lesson. At each level, contingencies were incorporated in the schedule to permit its adaption to individual students' responses, and to promote as much specificity as possible in students' responses prior to the research assistant's providing an external structure for their

LEVEL I
Understanding

Did you understand the material up to here?

Yes

No

Can you tell me what you were having trouble understanding?

Do you know why you were having trouble understanding ______________?

LEVEL II
Process

Did you understand why the teacher was ______________?

Yes

Why did he/she do that?

No

Do you think the teacher wanted you to be thinking in some (other) special way that would help you to understand (learn)?

Yes

What was that way of thinking?

No

Do you suppose that the teacher wanted you to be ________________?

...cycle through Levels III & IV for each positive response

(paying attention, recognizing, grouping, relating, remembering, inventing, making up, guessing, asking, repeating, using a special way to learn)

LEVEL III
Clarity

Was it clear to you that you were supposed to be thinking that way?

Yes

What made it so clear?

No

How might the teacher have made it clearer that you were supposed to be thinking ________________?

LEVEL IV
Relation of process to learning

Do you think that _________ helps you learn (understand, remember)?

Yes

How does it help you?

How do you learn by doing _____________?

No

Why not?

LEVEL V

Is there a better way for you to learn the material?

(ask a question, repeat the material to yourself, draw a picture, make up a little story, ...)

LEVEL VI

Did you know what would happen next?

How did you know that? What told you?

NOTES: Be sensitive to whether the current question has been answered already in the student's answers to previous questions. If so, skip the question.

Figure 1. Student Interview Schedule

comments, e.g., by providing alternatives from which they might choose or elaborate on.

The research assistants were trained to use this schedule in simulations prior to gathering data in the schools. These trials also produced several modifications to pilot versions of the final schedule that is displayed in Figure 1.

Procedures

Videotaping. Each teacher was assigned one research assistant for the duration of the study. After initial contact had been made with the volunteer teachers by the principal investigators, all subsequent scheduling for videotaping and interviewing was done jointly by the teacher and the research assistant. The research assistant was responsible for collecting all data and for acting as a liaison between the teacher and the project. A backup person was available in case one of the research assistants could not attend a taping session. This person, who also functioned as the administrative coordinator of the project, attended all training and other project meetings to insure comparability with the other research assistants. In only one instance was the backup person required to collect data.

Data collection began in the third week of February and was completed by mid-April for three teachers, and by early May for the other two, as shown in Figure 2. This variability was due to cancellation of sessions on the part of teachers as a result of conflicting school activities (e.g., field trips).

The procedure for data collection in the classroom was the same for all teachers. At the scheduled time, the research assistant went to the school with the videotape equipment and set it up either before class began in the morning or during recess. An information sheet provided procedural notes for setting up the video and audio equipment, and a checklist for its operation. Then she videotaped the lesson and took brief notes about instructional events that might be analyzed during the teacher interview, as described earlier. After the lesson, the research assistant took approximately ten minutes to move the playback equipment to another room and prepare for the teacher and student interviews.

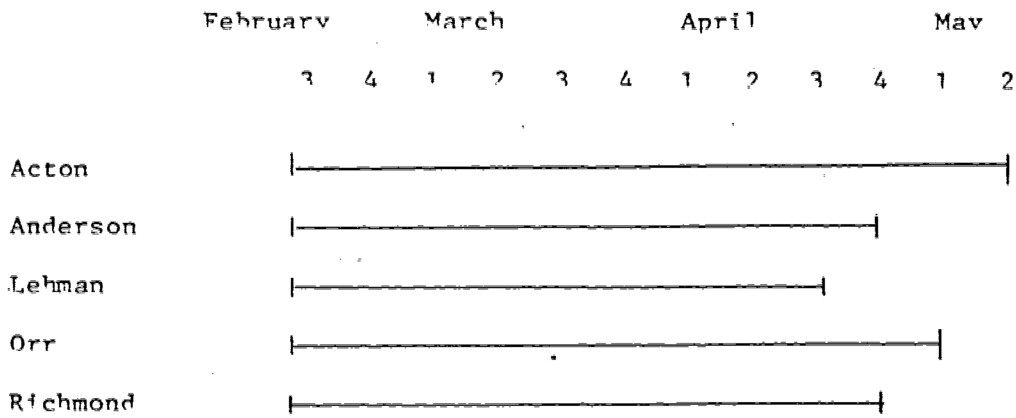

Figure 2. Length of involvement of each teacher, Study I

Teacher interview. The teacher viewed the videotape of the entire lesson. He was asked to stop the videotape at points where he intended the students to make particular cognitive responses to his teaching. They were prepared for this task prior to videotaping by a discussion of about a half hour in which we described examples of what we later define as instructional stimuli and intended cognitive responses (see Results section). One example we used in these discussions was a higher order question, where the teacher's choice of a particular syntactic construction is intended to signal students to engage in a complicated mental process such as analysis or synthesis. The research assistant also stopped the videotape at points she had noted in the log if the teacher passed over a large number or variety of salient incidents. It was made explicitly clear to the teacher that he should not invent an analysis of his intentions for students' cognitive mediational responses when he intended none.

During the interview, the audiotape recorder was turned off while the videotape was being viewed. Then, as either the teacher or research assistant stopped the videotape to analyze an instructional event, the audiotape recorder was turned on. Each time this occurred the videotape meter reading was announced by the research assistant or the teacher so that the audiotaped discussion later could be keyed to its respective segment on the videotape. These numbers, and brief notes that identified the intention being discussed, also were recorded on teacher interview log sheets so that the general features of an analysis could be examined without the use of playback equipment.

The surface features of the interchange between a teacher and the research assistant varied somewhat from the analysis of one incident to the next. In all cases, however, the goal of the analysis was to obtain from the teacher as precise a description as possible about the cognitive processes in which he intended students to be engaging in response to his teaching. Sometimes, the teacher reached this objective without prompting or requests for clarification by the research assistant. However, when probing was necessary, the research assistant followed the format

of the structured interview for students (see below). Most analyses took a form in between these two extremes of involvement of the research assistant.

The research assistant was careful to probe the teacher's descriptions in ways that solicited analyses from the teacher rather than asking him to verify whether one or another speculation offered by the research assistant was accurate. Only after probing of this sort proved insufficient for clearly describing aspects of cognitive processing the teacher intended students to use did the research assistant request clarification in ways that directly structured the teacher's description of his intentions for students' cognitive processing. This procedure was used to reduce the chances that the researchers' biases would influence the teachers' self-reports. However, we have no data that support the validity of this procedure, nor any reliability data which indicate that research assistants selected the same types of incidents for teachers to comment upon.

The teacher interviews ranged in length from 25-50 minutes. At the end of each interview the teacher was asked to identify several instances from the log about which he recommended the students be interviewed. In some cases (about 15% of the lessons), the teacher was unable to identify certain instances as being more appropriate than others. In these cases the research assistant selected incidents which she judged the teacher had stressed during the interview under the constraint that there be at least one selection from each of the beginning, middle, and conclusion of the lesson. Some of these selections may have been those selected by the research assistant for analysis.

Student interview. The research assistant spent approximately 10 minutes after the teacher interview preparing materials and planning for the student interview. Usually, three to four students were interviewed at a time, although occasionally slightly larger or smaller groups of students were involved in the interview due to factors not under the research assistant's control, such as school activities scheduled following the interview period. Both in the student interview log and on the

audiotape recordings, students' first names were identified. Also, the number of previous interviews each student had been involved in was logged. Whenever students were interviewed for the first time, they were informed of the purpose and procedures of the interview.

The research assistants were given a written statement which they paraphrased to introduce students to the interview procedures. It consisted of mutual introductions in the group and a brief description of what would be occurring during the interview, i.e., viewing the videotape and answering questions about how they were thinking during the lesson. Students then were shown preselected videotape segments, beginning a minute or more ahead of the actual event to be analyzed and usually going several seconds beyond the event so that they could re-experience the context in which the event occurred. Analysis of each incident was completed before the next videotape segment was shown.

During the first week of data collection, questions were directed to the students as a group. We discovered that further probing could not be done in a systematic fashion under these conditions. Thus, this procedure was altered for the remaining nine weeks. For the first event to be analyzed, one student was chosen randomly. This student responded to the entire set of interview questions about that instructional event before interviewing other students about that incident. However, to insure that other students' analyses of this event were not forgotten or lost, the others in the group were invited by the research assistant to add thoughts of their own after the set of questions comprising a major level in the interview schedule, (e.g., level II, process; see Figure 1) had been completed with the primary student. Thus, one target student's complete analysis of the instructional event was supplemented at each major level of the interview schedule with comments from other students in the group when they had different answers to the interview questions. This procedure allowed a full range of student responses to be obtained with the greatest possible efficiency.

When analysis of the first incident was complete, the research assistant randomly chose one of the remaining students and began the interview afresh following the viewing of the second incident. Whenever possible within time constraints for keeping the students out of class, each student in the group served as a target student for a unique instructional event. In those few interviews where relatively few instructional events were selected for analysis due to the nature of the lesson itself, or where a student was unable to answer any of the interview questions, more than one student was interviewed about a single instructional incident using the full interview schedule.

As with the teacher interviews, the playback of videotape segments was not recorded on the audiotapes, but references to videotape recorder meter numbers were announced as the analysis of each new segment was about to be recorded on the audiotape. Notes were written in the student interview log regarding any circumstantial conditions that were seen as unusual or of potential importance.

Transcription of interviews. To have data that could be easily manipulated during analysis, the audiotaped interviews were trancribed into written form. Perforated cards were selected for this purpose because they could be maintained in their original sequence or separated and rearranged, as the analysis proceeded. Teacher interviews were typed on white cards and student interviews on buff cards for easy identification.

To assure accurate, verbatim transcriptions of the interviews, research assistants dictated original interviews onto new audiotapes using standard instructions for resolving garbled or faint passages, for reproducing the interviews in a consistent format, and for inserting punctuation and identifying changes in speaker.

Data Analysis

Analysis of data represented by the videotaped lessons and transcriptions of teachers' and students' stimulated recall interviews about these lessons progressed through three stages. First, the research assistants' field notes and the transcripts of

teacher and student interviews were examined while we watched the videotape of each lesson. In addition to getting an overall impression of the lesson, we made notes and preliminary interpretations about the behavioral interactions the teacher and students displayed in the lessons and commented on in the interviews. We also studied the teachers' descriptions of the cognitive responses they intended their students to make alongside the cognitive activities students described in their analyses. The goal at this stage of the analysis was to be as inclusive as possible in representing the behavioral and cognitive features of the instructional milieu as these were represented on the videotape and retrospectively described by teachers and their students. Any reference by a teacher to intended cognitive responses for students, or by students to a thought process they used in response to some aspect of the lesson was noted in this first stage of the analysis.

With these notes, the videotape of each lesson was viewed again and analyzed to identify parallels and similarities among the instructional episodes, and the teachers' and students' descriptions of these episodes. The intent for this stage of the analysis was to produce a descriptive system for classifying instructional events in terms of behavioral interactions and the cognitive processing described by participants relative to that interaction. To allow comparisons of one instructional event with another, several alternative psychological terms were used to represent the cognitive responses that teachers and their students were describing.

The final stage of analysis involved classifying the products of stage two. Various schemes were considered with respect to: (1) their ability to capture the richness of the original lessons and reflect the descriptions provided by teachers and students; (2) the relative distinctiveness of categories within the system; (3) the linkage between the category labels and existing psychological theories; and (4) the ability to recreate the category system when index cards on which the descriptions were written were shuffled and resorted.

RESULTS

On each occasion where either a teacher or student remarked that a particular cognitive process was involved in learning from teaching, we assumed that something in the teaching environment signaled that the process was to be used. We labelled these signals instructional stimuli (IS). Each instructional stimulus was associated with at least one, and sometimes several, cognitive responses that a teacher or a student described in an interview. Conversely, teachers and students often associated several different instructional stimuli with one kind of cognitive response.

When teachers identified an instructional stimulus during their interviews, they were asked to describe how they expected students to respond cognitively to this cue. In other words, we asked teachers to describe how they tried to manage students' cognitive processing at key points in lessons by their deliberate use of instructional stimuli. For these occasions where teachers described instructional stimuli, we also asked the students whether the teacher wanted them to be thinking in a special way (see Figure 1, Level II). Thus, we collected students' interpretations of the cognitive processes they associated with the instructional stimuli their teacher had identified. From both points of view, then, a cognitive response associated with an instructional stimulus was one the teacher intended students to use, either as the teacher described the intention or as students perceived the teacher's intention. Therefore, we called cognitive responses to instructional stimuli intended cognitive responses (ICR). The pairing of an instructional stimulus with an intended cognitive response constituted what we label an instructional unit or an IS-ICR unit.

The cognitive mediational paradigm is based on a premise that students' cognitive processing mediates the effects that teaching has on students' learning. In our terminology, intended cognitive responses act as a filter, an amplifier, or a bridge between the instructional stimuli teachers use while they teach and students' memorial representations of the curriculum they construct during teaching. To reflect the central position of these psychological

events, we organized the interview data according to intended cognitive responses as they were described by students and by teachers.

Three major categories of intended cognitive responses to instructional stimuli were identified, each of which had one or two further levels of specification. The first category was orienting which referred to directing or redirecting students' cognitive processing to content or activities in a lesson. The second category was operating which entailed cognitive manipulations of information that were precursors to learning per se. The third category, consolidating, described intended cognitive responses that would enhance storage and retrieval of information or feelings. The full classification system describing intended cognitive responses that we developed is presented in Table 2.

Most intended cognitive responses had three features. One feature related to the type of goal toward which students were working or to a cognitive product that was to be achieved as a result of cognitive processing. Two illustrative types of goals are feelings and problem solving strategies. An example of a cognitive product would be an appraisal of one's work such as "I did very well at x." A second feature concerned the size of the goal or the amount of information to be processed. Some intended cognitive responses pertained to relatively simple instructional goals or a single item of information. Others required major shifts in memorial representations of content or the affective climate of the classroom, or large quantities of information to be manipulated. The third feature pertained to the immediacy of the intended cognitive response. Sometimes, students' cognitive processing was to occur at the same time as the instructional stimulus was used. Other instructional stimuli informed students about cognitive processing they were to use throughout a lesson. Yet another version of immediacy occurred when an instructional stimulus that cued cognitions appeared early in the lesson, but the cognitive processing itself was to take place later.

TABLE 2

Intended Cognitive Responses in Classroom Teaching

I. Orienting: Controlling the goals toward which students work.

A. Affective States:
- to promote engagement with lesson activities because these activities are enjoyable, they satisfy curiosity, or they increase arousal.

1. Engagement:
-- to elicit positive affective states as a side effect of engaging in an activity.

2. Arousal:
-- to increase or decrease anxiety, or other affective states.

B. Identifying Content:
- to orient toward specific content to be learned.

1. Macroscopic focus:
-- to create or to elicit a cognitive structure within which more specific lesson content could later be lodged.

2. Microscopic focus:
-- to attend to a specific item (as opposed to larger units of content) to be used either immediately or at a later point in the lesson.

3. Multiple microscopic items:
-- to attend to several specific items, as in microscopic focus.

C. Cueing Procedures:
- to cue procedures to follow in developing and acquiring content presented in the lesson. The teacher's intention has to be literally identified in the instructional stimulus.

1. Behavioral interactions:
-- to direct to an overt, behavioral interaction for controlling cognitive activity.

2. External materials:
-- to refer the student to material external to the student's memory that would facilitate cognitive processing.

3. Acquiring content:
-- to orient to cognitive processes applied in working memory.

4. Storing content in memory:
   -- to prepare to transfer the contents of working memory to long-term storage.

II. Operating: the way students think during instruction to achieve the intended products, such as recognition and comprehension.

A. Comparing:
   - to compare items of content to support acquisition of that content.

   1. Unspecified comparing:
      -- to engage in a general comparing procedure that would distinguish a cued item from other items of content; similar to microscopic orienting.

   2. Comparing codes:
      -- to facilitate students' acquisition and comprehension of content by having them compare the meanings that accompany different ways of coding that content.

   3. Comparing attributes:
      -- to compare attributes of two or more concepts.

   4. Using rules:
      -- to apply rules developed in the lesson to new instances; to compare the steps in a rule with aspects of a particular instance to which the rule was applied.

B. Generating:
   - to generate a new code for representing content; to generate new content using procedures or information provided in the lesson.

   1. Generating codes:
      -- to create alternate representations for information.

   2. Generating rules:
      -- to create a schema that provides a structure for generating a product.

C. Using Metacognition:
   - to think about cognitive processing and the results of that processing.

   1. Monitoring:
      -- to examine the cognitive procedures used to produce answers to questions or hypotheses.

   2. Appraising:
      -- to evaluate the state of knowledge and monitor the emotional response to that status.

III. Consolidating: practice designed to promote storage and retrievability of content. Students are expected to have already achieved comprehension of content prior to attempting consolidation.

Orienting

Almost all contemporary theories of cognition postulate that learners have goals toward which they direct their cognitive processing of information. This feature of the cognitive lives of students is clearly recognized by classroom teachers. Indeed, it is commonly accepted that one of the teacher's major objectives is to control the goals toward which students are working in lessons. When teachers used instructional stimuli to try to control the goals toward which students worked, or when students interpreted an instructional event as cuing them about goals, we categorized the intended cognitive responses they described as orienting.

Affective States

Two kinds of affective goals were seen to be end products of orienting. One goal was engagement. When this goal was approached, teachers and students described orienting to new lesson activities that would be enjoyable or that would satisfy curiosity. The other goal of orienting was an increase in students' level of arousal, presumably to some optimal level for learning. The cognitive product associated with both these goals was motivation to be actively involved in learning.

Engagement. Instructional stimuli intended to orient students toward activities were generally indirect or implicit. Often, teachers intended enjoyment to be a side effect of students' engagement in learning rather than being a result of what they would learn. Frequently, teachers had students give examples from personal experience to set the stage for engaging in subsequent tasks. At these times, teachers referred to activities that were "fun" as leading to engagement. For the most part, students did not notice instructional stimuli intended to orient them toward engaging in activities that would produce enjoyment. Most children said they focused only on the task posed by the teacher. Very few understood that the activity per se was designed to make them "enjoy" the lesson. For example, Richmond demonstrated his own enjoyment in a lesson by raising his hand in an animated manner to answer a question he had just posed. He

intended students to orient toward becoming personally involved in the lesson and thereby enjoying their participation in it. One student oriented to content instead of affect. He thought the teacher was going to answer the question. A second student oriented to procedures. He thought this was a signal for him to put up his own hand. Neither student viewed this instructional stimulus as a cue to orient to something affective about the lesson.

In at least one case, an instructional stimulus intended to orient students to enjoyment through engagement actually produced anxiety. Richmond started one lesson by stating a performance objective that would involve students in applying a rule to a real-life problem at the end of the lesson. One student responded as Richmond had intended by predicting this task would be easier because he could see how the lesson content related to his personal life. He was motivated to engage in learning content that would be presented later in the lesson.

> Researcher: Why do you think Mr. Richmond would want you to know what you were going to end up doing?
> Student: Well, it's easier to learn when you know what you're going to make -- sort of like making a model; so you just put the pieces together if you know what you're going to make first.

A second student reacted very differently. Because Richmond's objective appeared complicated, he panicked.

> Researcher: Was there anything, Mike, that was different for you?
> Mike: Yeah, he does it all the time, and I really don't like it because it kind of scares you at the start when he goes, "This is what you're gonna do, you're gonna learn how to measure a tree and how you do it outside," and I get scared because I don't think I'm gonna be able to do that ... So, you know, if you're gonna go outside and measure it, it

sounds so complicated to me, and I'm scared, and I just quiver a lot, because he -- well what happens if I won't be able to do that? So I kind of wish he doesn't do that right at the very start. You know, it's just kind of hard to do, because it does scare you a lot.

Researcher: I see, so it has a different effect on you, especially if what he says you're going to be able to do seems to you to be something you might not be able to do, and you think you'd be more relaxed and learn better if you didn't know.

Mike: Yeah and then he'd say, "How about you guys go out and measure a tree," and I say, "Oh, O.K., it's a challenge" -- but if he says it at the start, and then I don't pay attention, I'm trying to catch everything he says and I get confused.

In general, when these teachers used instructional stimuli to promote orienting to an activity that was intended to create positive affect as a cognitive product, the stimuli did not function as intended. This seems to reflect students' misunderstanding of the immediacy and the goal of these intended cognitive responses. Instead of looking forward to an activity with enjoyment, students usually focused on the content in the task or the task as a procedure. Rarely did students anticipate positive affect that might result from engaging in the task.

Arousal. Orienting students to engage in a lesson by increasing their arousal was cued most frequently by a teacher's use of the word "test." This was intended to orient students' attention immediately to lesson content or to procedures, and this effect was intended to endure throughout the lesson. Students' orienting was consistently cued when teachers gave this classic signal, and they almost always said this resulted in increased arousal. However, their understanding of how to go about learning material for a test was not enhanced by arousing them. On only one occasion did the instructional stimulus also help a student identify content: he said he should orient to "hard words" in upcoming parts of the lesson.

Neither teachers nor students reported that orienting to achieve arousal was likely to produce an effect that lasted beyond the immediate lesson. However, this cognitive mediational link was a strong one. Teachers and students were well aware of instructional stimuli that were intended to orient students to engagement by increasing arousal. Often, teachers did not intend students to make these links consciously. In some ways, the instructional stimuli were to evoke arousal like a classically conditioned response. By this, we mean that the teacher's instruction serves as a conditioned stimulus intended to evoke a relatively unconscious response from students regarding their current affective state. We speculate that teachers may be unintentionally establishing long-range affective traits in students as a by-product of their relatively immediate intentions for orienting instructional stimuli regarding students' motivation. That these rarely were mentioned can be construed as evidence that these considerations almost are taken for granted, like an involuntary response that follows the presentation of a conditioned stimulus. Here is one example Richmond provided when he was talking about his introduction to a science lesson.

> Researcher: So you wanted them to just pay attention?
> Richmond: I'm hoping that they tie this activity, this really happy, positive fun type activity in with anatomy, so that when they think "anatomy," they think that it was fun. They don't specifically know what happened, I guess, about it that was fun. All they know is we did something that was really a lot of fun.

Identifying Content

Intended cognitive responses where students oriented to content in the lesson were grouped in[illegible] three categories that varied according to the amount of the information. These three groups were macroscopic content, microscopic content, and multiple microscopic items. Orienting to content usually was viewed as a first step for other cognitive processes like compare or generate

that are discussed later. Orienting to focus on macroscopic content almost always was intended to endure throughout the lesson. In contrast, both kinds of microscopic focusing were intended to have immediate and short-lived effects on cognitive processing.

Macroscopic focus. The purpose of orienting students toward macroscopic content was to create a cognitive framework that they could use to organize more specific content that would be presented later in the lesson. The instructional stimuli that cued macroscopic focusing occurred at the beginning or very early in lessons. In some instances, the framework to be created was not directly related to the topic of instruction, as was often the case for Lehmann in his use of "anticipatory set." In his case, anticipatory set frequently was followed by a brief activity, e.g., having small groups of students (2-3) discuss with one another the defining characteristics of a topic introduced within the context of the anticipatory set he had provided. For instance, Lehmann began one lesson by describing a parking lot with its many parallel lines into which cars fit neatly. He intended this instructional stimulus would orient students to a rule to write neatly with each letter occupying a uniform space in a word, like cars parked between lines in a parking lot. Occasionally, these instructional stimuli resembled advance organizers (see Ausubel, Hanesian, & Novak, 1978). For instance, Richmond sometimes opened lessons with a question. This instructional stimulus was intended to orient students to information in the lesson, to help them encode it, and sometimes even retrieve it.

Students usually failed to interpret with accuracy instructional stimuli that cued orienting to macroscopic content. For example Acton started a language arts lesson by reading a paragraph, intending students to infer characteristics of stories that required inference making. However, all three students we interviewed perceived they were to analyze the content in the introductory story. Instead of orienting to global aspects of stories, thereby producing a framework for content to be

introduced subsequently, the students oriented to microscopic items in the story Acton had read.

In another case, Anderson intended macroscopic orienting to be students' response when he asked them to give illustrations of a particular concept and then selectively recorded some of their answers on the blackboard. He intended this to guide students in generating a cognitive structure for content to be introduced later in the lesson. The student interviewed was not aware of Anderson's intention for this instructional stimulus. Instead, he tried to memorize the microscopic items written on the board rather than creating a macroscopic framework.

When the purpose of an instructional stimulus was to orient students to macroscopic content, they usually tried instead to learn the microscopic content being presented at that time instead of creating a larger structure for later material. In other words, students often focused on the immediacy of the cognitive response, mistakenly thinking it should be applied to the content being presented instead of creating a framework for information that would emerge later and treating this cognitive product as an enduring guide to subsequent cognitive processing.

Microscopic focus. Instructional stimuli that cued students to orient to small bits of information required either an immediate cognitive response or one that endured throughout the lesson. When Acton's students were listing examples of a concept while he was writing them on the board, he abruptly asked the students for attributes of the concept. He intended to orient them to these microscopic pieces of information so they could use them to retrieve more examples. All three students interviewed found this confusing. They were busy wondering why the teacher changed topics so quickly instead of orienting as Acton had intended.

Question-and-answer exchanges were typical settings for microscopic orienting that was intended to last throughout a lesson. For example, Lehmann asked students early in a lesson to describe the concept of prime numbers. He wanted them to orient to these attributes because they would be used later with

fractions. One student failed to orient because he could not remember what a prime number was. Undaunted, he listened to another student's answer, believing it would be important for something later in the lesson. Another student described his cognitive response to this question as finding the definition of a prime number in his memory, figuring out what distinguished prime numbers from other kinds of numbers, and "locking it in your mind." He said he did this because he predicted this information would be needed later. Thus, by different means, both students achieved an appropriate cognitive product.

A second instance of microscopic focusing in question-and-answer exchanges occurred in Orr's class. He was rephrasing students' reasons for placing examples in one or another category, intending that students would orient to the specific criteria for performing this task. Both students we interviewed said this was exactly what they did. Moreover, one student said that Orr's rephrasings provided a model for his answers that he could use later in the lesson. This is evidence for an enduring effect at both microscopic and macroscopic levels.

Unlike orienting to macroscopic content, instructional stimuli that cued orienting toward microscopic content, either for immediate or for later use, were relatively well understood by students. Perhaps this contrast relates to observations that instructional stimuli that cued macroscopic content identification were introduced early in the lesson, before content had been developed. In contrast, some content and a context had usually evolved in the lesson by the time instructional stimuli cued microscopic content identification. This may have provided needed additional information that helped students perceive this intended cognitive response.

<u>Multiple microscopic items</u>. Some intended cognitive responses required students to identify more than one microscopic item of information and then operate simultaneously with these multiple items. Although this often entailed procedural cuing that we describe later, several incidents are included here because they emphasized microscopic content identification. This

overlapping among cateogires illustrates the difficulty of separating some classes of intended cognitive responses.

The instructional stimuli that cued orienting to multiple microscopic items sometimes were the simplest we observed. For example, both Orr and Acton used obvious cues on the chalkboard (e.g., underlining or using colored chalk), intending students to identify a concept and its attributes, and to compare this concept with some other concept. Students consistently attended to these cues, but they had less success matching their cognitive responses to the teachers' intentions. In a lesson on fractions in which Orr used colored chalk to focus students' cognitive processing on multiple items, one student indicated that he focused on a single marked item without referring to any of the others that Orr wanted him to compare to the identified one. Two other students interviewed also noticed Orr's cues, but didn't know how to think differently about the concepts they referred to.

In Acton's class, underlining a word on the chalkboard to indicate it as an instance of a concept was intended to cue microscopic content identification that would be followed by comparing. The student interviewed about this incident responded perfectly. She noticed the visual cue, oriented to the particular attributes of this concept, and then proceeded to extend cognitive processing to analyze why one item was underlined and other items were not.

What may characterize these events as a group is that students often could not associate a specific cognitive response with the instructional stimulus because this pairing either had not been practiced enough or was too inconsistent. Our speculation about this is based on an occasion in which an overlearned response was cued by a visual stimulus. Specifically, Acton was using well-learned symbols to indicate how students should pronounce words in a language arts lesson. The student interviewed about this incident indicated he knew the meaning of these visual cues and also said he understood that these cues referred to specific concepts that would be involved in a subsequent cognitive task. Thus, instructional stimuli that cue

microscopic content identification followed by subsequent cognitive processing may not be well understood by students unless both the concepts being identified and the subsequent cognitive processing implied by the various instructional stimuli are clear or well-practiced responses.

Cuing Procedures

There were four procedures toward which students were oriented by instructional stimuli: behavioral interaction, use of external materials, use of a cognitive process to develop content (what we develop later as several types of cognitive operations on content), and storing material in memory (which we treat subsequently as consolidating). The data we classified as procedural cueing are somewhat difficult to distinguish from orienting toward microscopic content followed by another cognitive process (like comparing), and from fairly direct injunctions from the teacher for students to engage in cognitive processes such as comparing. This also reflects the overlapping and correlated nature of these phenomena in their natural environment. The rule we used to place incidents in the category of procedural cuing was whether the teacher's intention was literally identified in the instructional stimulus. If this intention was implicit, the incident was categorized as cuing procedures. If the intention was explicit, the event was placed in the domain of cognitive operations, described later.

Behavioral interactions. This category is unique in that an overt, behavioral interaction was used to promote cognitive activity. For example, a student who had just given a wrong answer was told to listen to a second student's answer. This was an instructional stimulus by which the teacher intended to orient the first student to be ready to paraphrse the second's correct answer. A second incident, also in Lehmann's class, was one in which he had students engage in a small group discussion, the topic of which was the nature of small group discussion skills. The intention was that this activity would orient students to more efficient ways for developing lesson content in small groups. A

student interviewed on this point reflected this intention precisely. A third incident in this subcategory occurred in Anderson's class when he said, "I'm going to check with you to make sure you've learned what I've said." He intended to orient students to a procedure that would be followed throughout the lesson. This enduring intended response was confirmed precisely by one student we interviewed, and more vaguely though still accurately by a second.

These instructional stimuli were clearly understood by the students. In fact, students had a well-developed understanding of general behavioral interaction patterns in their classrooms. For example, when asked how they knew what the teacher would do next, students frequently replied, "Because he always does it that way." Intended cognitive responses were interpreted accurately by students when the link between an instructional stimulus and the intended cognitive response had been practiced frequently.

External materials. Teachers often used instructional stimuli to orient students to places, usually the chalkboard or the students' notebooks, where content or operations for manipulating content were presented. Teachers reasoned this would ease the burden of cognitive processing for students. They wanted students to use this information either later in the lesson or in reviewing lesson content for a subsequent purpose, such as a test or a laboratory exercise.

One example ocurred when Richmond dictated information for students to record in their notebooks. He intended students to refer back to their notes later while learning other material. He said that students should find this activity useful in helping them transfer this material to new tasks. A student we interviewed saw a clear advantage to taking notes, particularly when the teacher told him what to put in his notebook. This strategy is efficient, he said, since he would not have to remember all of the content but only the part of it not included in his notes. However, the student made no reference to the teacher's intention to enhance transfer by this procedure.

A second example was illustrated by Anderson's review question, "What is the important thing you have learned?" He intended students to orient to information on the chalkboard. One student's view of the intended cognitive response was like Anderson's: she made explicit use of the external material in order to answer the question and, furthermore, sorted it into superordinate and subordinate categories. A second student perceived the intended cognitive response as "(I) think in my head ... or say it quietly to myself." Though the idea of reviewing material was apparent in this second student's perception, he made no use of information on the board. The cue to orient to a procedure failed, but the intended operation materialized nonethless. Whether the cognitive product resulting from this response was as successful as it might have been if the student also had oriented to the chalkboard, as was intended by the teacher, is moot.

Acquiring content. In information processing terms, some instructional stimuli were intended to orient students to use particular cognitive operations to work on information in their memories. Sometimes, teachers attempted to orient students away from acquiring specific content to applying a rule that would integrate the content. The instructional stimuli cuing this procedure were direct and explicit, such as writing the rule on the chalkboard. Students' perceptions of the intended cognitive responses cued by these instructional stimuli matched their teachers' almost exactly. For example, in Acton's class, an outline on the chalkboard signaled to students that they were to reorient their cognitive operations on materials being dealt with. Whereas in the preceding part of the lesson they had been generating ideas about a story they had read, the students now were asked to apply rules to describe how an author might end a story like this one. Another occasion where the teacher's intention was to reorient students' cognitive operations in working memory occurred in Anderson's class when he wrote a rule on the chalkboard. Here, students said things such as: "(I should) match them together, what's on the chalkboard and what's

in my head." Overall, students' views of the intended cognitive response were congruent with Acton's.

Commonly, teachers asked students to begin to monitor the way they were learning as a cue to orient their thinking strategies in working memory. For example, Lehmann often imposed specific time limits on tasks for students. He wanted students to know that they should monitor their cognitive processing (e.g., apply a rule) to increase efficiency. ' student interviewed about this instructional stimulus had a g ,,, ral sense of this message, though his descriptions of the intended cognitive response were very vague (e.g., "Don't fool around").

The immediacy of these intended cognitive responses was clearly understood by students, possibly because of the distinctiveness of the teachers' instructional stimuli. Students apparently had overlearned responses to most of these kinds of instructional stimuli. This probably accounts for the strong correspondence between teachers' and students' reports of the intended cognitive responses in this category.

Storing content in memory. Orienting students to transfer the contents of working memory into long-term memory was common. In two different lessons, Orr used cues that appeared to be opposites even though he described the same intended cognitive response for both. In one lesson he asked, "Do you know what we will do next?" In another, he said, "Let's back up a bit." Both times, he intended students to switch from acquiring content to reviewing prior content and storing it. The students understood and made cognitive responses congruent with the teacher's intentions. For example, after the instruction to back up, one student reported that she first re-organized the information she was thinking about, then recalled all the previous items in the order they had been presented so she could analyze them in terms of her new definition for the concept. Finally she used other students' answers to practice recalling the rule she generated. A second student gave an almost identical description.

Not all instructional stimuli teachers used to cue this cognitive response were perceived accurately by students.

Students occasionally made intended cognitive responses even when the teacher did not clearly understand what he intended. Acton made a statement in one lesson: "Think really hard." He described this as "It's just an expression ... It means think hard ... I don't know." Both students interviewed about this incident interpreted this as an instructional stimulus. One student believed the intended cognitive response was to orient to content that would appear on a test. He said: "In a way it's like a threat ... Start thinking or else you're going to fail." The second student also felt anxious and attempted to lessen his anxiety by reviewing what they had just been studying so he could understand "what he wants us to do." Thus, even if teachers do not intend some events to be instructional stimuli, students nonetheless may perceive and respond to events using cognitions they think are intended by the teacher. The differences between the two students interviewed about this event also supports our speculation that variability in students' on-the-spot cognitive responses to instructional stimuli may result in different effects of teaching practices on different students' learning. Finally, this incident also illustrates the multidimensional nature of instructional stimuli as signals for intended cognitive responses. Not only was this instructional stimulus associated with orienting students to a procedure, but it also is clearly linked to orienting by arousal.

Operating

In the preceding section, there were examples of orienting responses which were to be followed by various kinds of cognitive operations by students. Most labels for acts of thinking (e.g., attention, comprehension) describe a product of cognitive operations rather than operations per se. Our objective, however, was to describe the operations that students use in response to instructional stimuli. Thus, in this section, we avoid labels like "attention" or "comprehension" that describe cognitive products.

A survey of about thirty books about human learning and cognitive processing suggested that there are three basic

operations that learners use in instructional settings: comparing, generating, and re-examining. These three operations may not be completely separable from one another. Achieving a cognitive product such as comprehension probably entails applying all three processes in varying amounts and sequences. Nonetheless, we found this three-part scheme useful for analyzing teachers' and students' views of intended cognitive responses to instructional stimuli.

Like orienting, operating also can be described in terms of size, i.e., the amount of information being operated on, and the immediacy with which the operations should be carried out. Though it is tempting to create a table that matches each kind of operating with the different levels of focus and immediacy, we believe the data and the classroom phenomena that they represent do not justify this sort of factorial arrangement. Thus, the following analysis reflects naturally overlapping and correlated features of students' cognitive processing during instruction.

Comparing

When the intended cognitive response involved comparing aspects of content to promote learning, it always involved an immediate cognitive operation. There were four basic kinds of intended cognitive responses that involved comparing.

Unspecified comparing. This intended cognitive response is similar to orienting toward microscopic content. A cue on the chalkboard always served as the instructional stimulus. While teachers did not elaborate about the nature of comparing they intended students to undertake, hence the modifier "unspecified," it was clear that students were to go beyond mere orienting to distinguish the cued item from others by using some cognitive operation.

Students' views of intended cognitive responses to these instructional stimuli were varied. When Orr circled items on the chalkboard, two students said they were confused about the comparing Orr intended. Further probing revealed that one student had mastered the content under discussion but was not sure about

how the teacher's instructional stimulus was to be interpreted. The second student clearly noticed the instructional stimulus and indicated that he knew some sort of cognitive operation was to be invoked, but he did not know exactly what process should be called upon. Beyond these two kinds of confusion, a third student thought that he was to compare specifically the attributes of the item which Orr circled with the attributes of the items surrounding it on the chalkboard. Despite the lack of specificity in Orr's intentions, this third student proceeded to engage in specific cognitive processing. This is another instance of a student engaging in processing not perfectly matched to processing the teacher intended. (In this case, the teacher was not clear about what was intended.)

This incident showed that an instructional stimulus can communicate different intentions to students depending on their state of knowledge of lesson content. The first student interviewed understood the material. She did not perceive any need to make comparisons. The second student, who had not mastered the lesson content, was unsure about the intended cognitive response to the instructional stimulus, though he perceived that something was intended. However, the third student, who was consolidating content from the lesson, matched Orr's intention. He was at the appropriate stage of learning to profit from a cognitive response that was partly his own, while being generally congruent with the teacher's vague intention.

Comparing codes. The teachers we worked with often were sensitive to how well students understood material, and how much of the lesson's material was understood. One way they showed this concern was by having students compare the meanings that accompanied different ways of coding information. Quite a variety of instructional stimuli cued this intended cognitive response. For example, in a lesson on anatomy, Richmond briefly described anatomical structures, then had students touch parts of their bodies. He expected them to be comparing attributes of the abstract concepts, presented in the lesson, with tactile sensations illustrating those attributes.

In addition to having students engage in inductive activities to compare codes, Anderson also engaged in activities himself to communicate this intention. On one occasion, he strutted emphatically across the front of the classroom, intending to signal that students should compare his physical action with aspects of a question involving a previously presented abstract concept. A student interviewed about this occasion clearly understood the intention communicated by this instructional stimulus.

> Researcher: Did you understand why he was doing that?
> Larry: Because he was walking and trying to prove something.
> Researcher: He was walking, and you said, when he asked you for an example, you said "walk" didn't you? Do you remember how you came up with that example?
> Larry: He was walking across the room, back and forth.
> Researcher: So why do you think he was doing that?
> Larry: To give us an answer. To give us a hint.

In another example, Acton wrote equivalent forms of fractions (e.g., 1/4 = .25) close to one another on the chalkboard. He intended the proximity of the two codes to cue students to use the two different codes for representing numbers. Both students interviewed perceived this comparing process, describing it first as using a familiar code to identify critical attributes.

> Researcher: Can you tell me how you were thinking, or how that helped you think about it?
> Kathy: He went back to the fraction, you know, and without changing to decimals it makes it easier to think more -- it's much easier to concentrate on it.

The teachers used obvious cues to inform students that they should compare codes. Students rarely misperceived these instructional stimuli and, as long as they had the knowledge to compare, had little trouble doing the comparing.

Comparing attributes. The operation of comparing codes is similar to that for comparing attributes that define concepts. Whereas the previous category pertained to attributes for a single concept, this category involves comparing attributes across two or more concepts.

When teachers cued this process, they intended students to focus immediately on content at a microscopic level. For example, in a grammar lesson, Acton read sentences to students after instructing them to "listen." The students' task was to identify adjectives and adverbs and compare their descriptive qualities in preparation for a later writing assignment. The students interviewed perceived this cognitive response as Acton intended. One student said he had a prototype of each concept in his mind. Upon hearing a particular descriptive word, he tried to examine that word in terms of the attributes of his prototypes.

> Researcher: Can you tell me what you were thinking as he was going through that?
>
> Bob: Well, right away from the first couple of lines of the paragraph I knew the kinds of words that I was going to use and so I was pretty well ready to write from there.
>
> Researcher: So that his reading that right at that point helped you to identify the kind of words you might need later?
>
> Bob: Yeah.
>
> Researcher: I'm trying to find out if you were just concentrating on the words he was saying and identifying, "that's a good word and that's a good word," or whether you actually were picturing what the whole example was about?
>
> Bob: It was kind of a little bit of both and when I heard certain words, right there I knew that those were some of the words that I wanted to hear. But I like to take the whole thing in general and look back over it after I have heard it and then I think about it just before I write down what I'm going to write.

Researcher: Was it clear to you that you were supposed to be listening and picking out those words?

Bob: I don't know if that was what we were exactly supposed to do. I just thought that he just did that to give us a good idea to start us with what we are going to try to do with the words.

Researcher: Does that help you to learn, if he reads an example to you?

Bob: Yeah, it does for me because, like I said before, I can pick up the things that I want to use and put them into my own words and examples.

In another lesson Acton had very recently been asking students to describe why they chose to discuss certain items as opposed to others. Then, he abruptly asked students why they had not chosen to examine other items. By asking students to focus on the criteria they used in deciding not to discuss a particular item, he intended them to compare these criteria with those they had used to select other items so the latter would become clearer. We interviewed three students about this event. None of them described cognitive operations very congruent with Acton's intentions. Two students had no idea why the teacher had changed the focus of the activity, so they decided to treat Acton's remark as a discrete question rather than one related to the previous work they had been doing. They had little idea what was intended by the current activity, nor could they predict what would be happening next in the lesson. The third student, however, perceived that Acton had an intention for some cognitive processing, but he understood his cognitive task to be one of isolating a separate set of attributes that would appear in his and other students' answers to the teacher's new question. However, he did not perceive that he was to make direct comparisons between the set of attributes currently being developed and those that had been developed earlier. Thus, he was unable to relate the new activity to the previous task, and thereby failed to highlight the basis for classifying items.

The teachers often read material to students. One such occasion occurred when Anderson read sentences in which "to," "too," and "two" had been blanked out. He intended students to compare the semantic attributes of various sentences and use them as a basis for deciding which form of the word to use in filling in a blank. One student we interviewed about this was candid in admitting that, "I knew what we were going to be doing ... I got kind of confused ... 'cause I was kind of fooling around." However, in defence of this student, we note that he also said he did not need to pay attention at this particular point in the lesson because he knew the correct uses of these words. Here, again, is an indication that a student's perception of the cognitive processing a teacher intends can be conditional upon the student's mastery of content under consideration.

Finally, one incident belonging in this category illustrates the earlier speculation that students can generate their own interpretations of instructional stimuli when these intentions are absent or vague in the teacher's description. On this occasion, Acton was asking the students to relate sentences they had written to a picture that had served as a focus for the writing activity. His description was, "I don't know how they'll be thinking about it here, just be putting it all together." In response to probing, he suggested that the students should have been comparing attributes of the material they previously had written with the concepts presented in his questions. The student interviewed about this incident had quite a different perception of the cognitive activity intended by the teacher. He attempted to store in memory other students' answers because he believed the teacher's questions would be asked on a later test. Because he was occupied with trying to remember the other students' answers, he said little else was being done. Thus, the student was clear but incongruent with the teacher on this specific occasion when the teacher was not very certain about how students should be operating on content.

Using rules. When teachers asked students to apply rules to new examples, we categorized these events as comparing because

teachers and students indicated that students were to compare the steps of a rule with characteristics of the particular example to which the rule was being applied. For example, in a lesson on fractions, Lehmann developed a step-by-step example on the chalkboard, talking about each step as he went and emphasizing what would follow each step. He wanted students to predict the next step based on the preceding step. The student interviewed said he was comparing the steps of the rule he had learned for working with fractions with those being illustrated on the board. His prediction for the next step involved comparing what needed to be done to solve the problem with the information provided up to that point.

Not all students accurately perceived teachers' intended cognitive responses involving rules. Following the presentation of a complicated rule, Orr asked several questions of students, none of which produced answers. He then said, "That's tricky; let me back up." He intended that students would compare attributes of the rule he had presented earlier with the material that was going to be re-examined. One student indicated she understood that previous answers were wrong, and that she needed to understand the steps in the rule. However, rather than analyzing earlier examples in terms of the rule, she perceived that the intended cognitive response was to orient to the next answer to seek attributes of it that could to be compared with the rule. A second student's intended cognitive response was congruent with Orr's. He reviewed the rule and tried to predict examples that might fit this rule.

Generating

Intended cognitive responses for generating new content involved either generating a new way to represent content, or generating new content per se based on procedures or information provided in the lesson. The former were rare, but there were a number of incidents in the second category.

Generating codes. Some teachers believed students would be able to acquire or retrieve information more easily if they had more than one way of cognitively representing that information.

Unlike the comparing codes response described before, where both coding schemes for comparison were already available, here students were to create a new code.

A prototypic incident was where students were to create a mental image. Before defining "concave," Orr asked, "Where do bears sleep?" He thought students would create an image of a cave that would help them remember the definition. Both students we interviewed noticed this odd question and understood that it was supposed to provide a device for learning the definition. However, neither perceived that this device was to be an image. When Orr clarified this by describing the image he intended students to generate, both students then said they recognized this as the image that Orr intended them to create originally.

Generating rules. The teachers frequently tried to cue students to generate a rule that could be applied to content. This process closely resembled hypothesis generation like that used by scientists. The teachers always assumed that students understood how to use such rules. Nowhere in their classroom behavior nor in their descriptions of their intentions for students' cognitive processing did they indicate that students might not have been aware of the form for a definition, an hypothesis, or a rule.

As one example of generating rules, Anderson expected students to create a rule about transitive and intransitive verbs as he categorized examples of those two verbs on the chalkboard. The student interviewed about this said, "He was trying to get us to fit the right ones," a weak concurrence with the teacher's intention.

In another example, following a discussion about stories and how to write them, Orr asked questions about how to end a story. He intended students to create endings by first devising a rule about the story's plot and then applying this rule. The student interviewed indicated he was baffled by the teacher's instructional stimulus. Although he understood that the story was to be ended, he had no idea how to go about it.

On two occasions where the intended cognitive response was to generate a rule, the instructional stimulus was something written on the chalkboard or an incomplete sentence that would define a concept when completed. We considered these definitions to be rules because they could be used to classify instances and non-instances of concepts. Students' interpretations of their teacher's instructional stimuli depended on whether they knew when processing was complete. In Orr's class, one student said he did not understand why the teacher had written an incomplete sentence on the board. A second student told us she believed she was to guess what might go in the blank, but she did this without referring to previous lesson content or to a rule that could make this task less hit-and-miss. A third student, however, described the same intended cognitive response as the teacher, namely, to complete the sentence and test its adequacy against information presented earlier in the lesson.

Students' interpretations of instructional stimuli that cued rule generation were clearly related to their understanding of the lesson content preceding these events. Students who were at the point where a rule usefully summarized what they had learned seemed to achieve fairly high congruence with the teacher's intended cognitive response. When their understanding was incomplete, students could not evaluate how successful they were at completing the task. A typical strategy was to wait for other students to complete the task. Their reluctance to ask questions about the meaning of these instructional stimuli appeared to be not so much an attempt to avoid displaying their ignorance as it was reasonable strategy for getting the needed information.

Teachers also intended students to create hypotheses. This differs slightly from rule generation because a rule could be used immediately and systematically following its generation whereas an hypothesis would become the object of further testing.

In a science lesson, Richmond asked students to guess the relationship between two variables by generating hypotheses. Students described a variety of cognitive responses to this instructional stimulus. One student interpreted the instructional

stimulus to mean that he should "think harder," which meant "to think over and try to remember" some of the preceding content. When asked to elaborate, the student said that he was to respond with a guess, but he did not know the criteria that determined what qualified as a good guess. A second student's interpretation of the instructional stimulus went beyond the cognitive response Richmond intended. This student not only created an hypothesis about the relationship between the two variables, but believed he would be expected to "tell why" this hypothesis was likely to be true.

There was considerable variability in students' interpretations of instructional stimuli that teachers intended as cues for students to generate an hypothesis. Part of this divergence seemed to arise because students did not know when their hypothesizing was adequate. Those students who could judge the adequacy of their thoughts also seemed to understand how to generate hypotheses. Occasionally, these students even went beyond the teacher's intended cognitive response to question the validity of the hypotheses they had generated. When their understanding was incomplete or absent, the task was relatively opaque with respect to the cognitive operations the teacher intended students to engage. For students in this situation, a typical strategy was to sit tight and wait for other students to complete the task. It appeared that their reluctance to interrupt the lesson with questions about the teacher's intention was not so much an attempt to avoid incriminating themselves as it was reasonable strategy for getting the right answer. On the other hand, students who were at the point of acquisition where a rule could be useful in summarizing what they had learned seemed to achieve fairly high congruence with the teacher's intention. Thus, instructional stimuli that cued students to generate either rules or hypotheses varied in clarity as a function of two factors: whether the student's mastery of content was sufficient to complete the task; and, whether the student could judge when generating was sufficient.

Using Metacognition

Metacognition refers to thinking about the processes and outcomes of thinking. Several forms of metacognition appeared in response to instructional stimuli, all of which were macroscopic in focus. There were two functions for metacognitive responses. One was to monitor cognitive operations on content to insure that the cognitive strategy being used was the one intended by the teacher. The second function involved something students were to do upon completion of monitoring. One some occasions, they were to provide feedback to themselves about learning. On others, they were to generate for themselves a motivational consequence based on their performance.

Monitoring. Teachers frequently intended students to examine the cognitive procedures they were using at particular points during lessons. For instance, in a math lesson, Richmond summarized a complicated principle by drawing a diagram of the principle. He intended that students would monitor the extent to which their sequence for applying parts of the principle matched his diagram. Specifically, students were to examine whether a preceding part of the principle cued a subsequent part. The students interviewed could not do this because they had not mastered the components of the principle. Thus, they could not use them as cues for subsequent components. Although the students had a general idea that they were to be monitoring their thinking, they treated each element as an individual entity. When asked what he did when he didn't know a component, one student said, "I erase the whole thing and start over." Monitoring whether he was applying components of the principle in sequence while trying to acquire the next component was beyond his existing state of knowledge and the capacity of his working memory.

In a somewhat similar situation, Acton was summarizing content in closing a recitation lesson. His intention was that students would monitor their seat-work activities using the schema he modeled while presenting the summary. Though this instructional stimulus and the one just described for Richmond both occurred at the very end of a lesson, they differed with respect to immediacy: Richmond intended students to engage a

monitoring operation simultaneously with his presentation while Acton intended students to apply this procedure at a later time.

Intended metacognitive responses were applied in the service of acquiring material as well as reviewing or applying it. Typically, this took the form of students answering example problems. For instance, when Acton did exercises illustrating how to change fractions to decimals, he intended students would monitor their use of rules for answering the exercises and compare their procedures with the one he was modeling. The three students interviewed all perceived this intended cognitive response. One student put it quite eloquently: "He's up there for a reason -- so we can think about what our answer is and how we do it." In a somewhat different vein, by providing incorrect examples, Anderson wanted students to monitor the rules they were using to classify instances of a concept. He believed that by monitoring, students would discover a relatively common mistake and remember that it was to be avoided. The student interviewed about this incident also described cognitive activity in consonance with the teacher's intention.

Only rarely were students baffled by an instructional stimulus cuing them to monitor. The few failures were due to a lack of content knowledge. When the content was not well learned, students focused on acquiring the content, not monitoring. This seems sensible since it is difficult, if not impossible, to acquire content and monitor its manipulation simultaneously.

Appraising. Teachers were aware that students judge what they know and that they react positively or negtively to those judgments. Teachers sometimes used instructional stimuli to cue students to do this. At the end of a lesson, Lehmann asked a student to summarize what had been covered. He intended that the student doing the summary would monitor the extent to which he had mastered the content. Where the student judged that mastery was inadequate, there was an unstated requirement that the student would request help from the teacher or peers. In Lehmann's view, this procedure provided confirmation of mastery and an incentive for monitoring. All the students interviewed about this

instructional stimulus indicated that they understood the intended cognitive response regarding both knowledge and affect.

An interesting illustration of Orr's intention for students to engage in appraising following metacognitive activities occurred in a recitation setting. He had been asking a series of questions about a complicated principle, none of which the students could answer to his satisfaction. When Orr was asking these questions, he expected students to be monitoring the cognitive procedures they were using to produce answers (as well as using rules; see the section on this operation). Upon seeing that this was impossible because students had not acquired the content whose manipulation was to be metacognitively monitored, he said, "That's tricky; let me back up." His intention concerning the first part of this phrase was that students should examine their feelings about their inability to provide answers and attribute failure to the difficulty of the task and his poor judgment about the extent to which necessary content had been acquired.

The two students interviewed about this event had quite different understandings of Orr's intentions concerning appraisal operations. Both indicated clearly that they knew something was wrong when no one, including themselves, could produce answers to the questions. Neither indicated that they were engaging in metacognitive monitoring, which confirms our prior speculation about the inappropriateness of this task when content has not been acquired. One student understood Orr's statement to mean that he should stop trying to answer the questions and note Orr's next comments as the answers to the previously unanswered questions. The second student understood Orr's statement about changing the direction of the lesson as a cue to review the previously unsatisfactory answers for common elements. Thus, the teacher's intention that students' attribute their failure to task difficulty was not perceived by these two students.

Consolidating

The third major class of intended cognitive responses pertained to storing and retrieving information. These activities

typically are operationalized in activities that involve practice. The feature distinguishing consolidating from the preceding cognitive processes (i.e., comparing, generating, or metacognitively examining cognitive procedures) is that teachers expected that students must already have comprehended lesson content prior to consolidating it. In contrast, the cognitive operations discussed earlier were intended to promote comprehension per se.

There were two components of consolidating. One was simple rehearsal. A second was to practice procedures for retrieving information based on cues similar to those presented in the lesson. This produced either verbatim recall or transfer depending on whether cues were the same as, or different from, information presented in the lesson.

Teachers believed that rehearsing and practicing retrieval strategies would increase the distinctiveness of information, the probability that information would be retained, and students' ability to retrieve information if they were given cues. This description is similar to contemporary theories of storage and retrieval (e.g., Gagné, 1978; Mayer, 1979).

Consolidating operations occurred throughout lessons. Lehmann began one lesson by having students discuss among themselves procedures they would follow while working later in small groups. In addition to orienting students to lesson procedures, he intended that they would recall and rehearse the skills involved in small group work. Students interviewed about this event perceived this same cognitive operation.

Instructional stimuli cuing consolidating operations also were used as teachers developed the content of lessons. In a vocabulary lesson, Acton told the students to "just jump right in" after presenting the third example of a concept, intending students to practice retrieving and applying a previously learned rule. One student clearly matched this intended cognitive response. He knew what the rule was and said he was going to apply it again, "because we already know the process to get it." A second student, though he knew what was intended, did not make

the cognitive response because he did not understand the rule to be rehearsed.

As might be expected, students' perceptions of consolidating responses to instructional stimuli were not always all that the teacher might have wanted. Lehmann frequently asked students to check with partners about their understanding, sometimes as often as ten times during a 45-minute lesson. His intention in one arithmetic lesson was for students to retrieve rules and rehearse the sequence involved in applying the parts of them. A student we interviewed about this event understood Lehmann's intention and felt competent in executing the consolidating processes, adding that his first step was to select the easiest aspect of the rule to retrieve and rehearse, thereby making the task easier.

The majority of instructional stimuli that cued consolidating responses occurred at the end of lessons or as transitions from recitation to seatwork. In most situations, students understood that they were to consolidate, though sometimes they were confused about whether to rehearse or to practice retrieval given cues. For example, in one lesson Anderson spent extra time reviewing, intending that students would practice their retrieval strategies and highlight the differences among items. Though both students interviewed understood that consolidating was the objective, neither knew how to proceed because they did not know the specific processes to use to consolidate lesson content. This was not always the case, however. By presenting steps in a rule very slowly, Richmond expected students to be rehearsing each step and improving their recall of each step by linking it with previous parts of the rule. The student interviewed about this event indicated clearly that he viewed his task as practicing each step and relating it to the preceding step.

On one occasion, Acton told his students to use a rule over the weekend that they had developed in class. He intended students would consolidate their strategy for retrieving the parts of the rule and rehearse its components. One student remarked that she understood the teacher's intention, but she would not do it. In her words: "Well, I would just look at it and then try

and remember it for the weekend. I don't think I would do anything like trying to convert it or anything -- I would just keep it in mind." A second student had the same reaction. Thus, unless the teacher can create tasks that control whether students make consolidating responses, students may not make those responses even when they know what to do.

DISCUSSION

We constructed a complicated system to represent the range of intentions that teachers had for students' cognitive processing of content in lessons and to reflect students' views of the cognitive processes they perceived were intended by their teacher. This system contains three major categories: orienting, operating, and consolidating. Orienting and operating each were further divided into three sub-categories, and these, in turn, were divided further into a total of eight tertiary classifications (see Table 2). Clearly, both teachers and students in our study reported that teachers used instructional stimuli to try to manage a large variety of cognitive operations they believed mediated students' learning.

The findings about teachers' and students' views of students' cognitive processes for classroom learning have several important implications. First, there was a noticeable lack of one-to-one correspondence between cognitive responses teachers associated with instructional stimuli and the cognitive operations that these stimuli cued for students. This suggests that research using only classroom observation schedules as the basis for generating explanations about how teaching affects learning may distort the ways in which these phenomena interact. Whereas observation schedules require that instructional events be defined as distinct if they differ in behavioral structure, teachers and students may treat them as functionally equivalent in terms of intended cognitive responses. And, while some instructional stimuli may appear superficially similar at the behavioral level, they may produce quite different cognitive responses by students.

Thus, any analysis of classroom instruction that does not account for the dynamic nature of the cognitive milieu will probably distort the phenomena of interest. In practical terms, if teachers do not communicate clearly the relationships between what they are teaching, how they are teaching, and how students should be thinking, students' learning may not be optimal.

A second important finding is that the teachers and students we interviewed did not presume that all teaching events lead directly to learning. Teachers used instructional stimuli to prepare, orient, guide, and otherwise involve students with information before students were to learn per se, and students perceived similar intended cognitive responses to instructional stimuli. Insofar as we know, such "supportive" or "preparatory" teaching events have been investigated rarely in classroom-based research on teaching. Separating supportive or preparatory instructional events from those intended to produce learning directly may be a distinction to be pursued in future research. This distinction also may prove to be useful in classroom instruction designed to help students become more deliberate in their use of different cognitive strategies to achieve various goals.

Turning to the findings about students' perceptions of instructional stimuli, five salient generalizations emerged from our analyses of students' understanding of the intentions teachers tried to communicate during instruction by using instructional stimuli. First, there was a strong probability that students would not accurately perceive a teacher's instructional stimuli that signaled students to use orienting operations. This occurred especially when teachers attempted to promote affective states in students. To a lesser extent, mismatches of this kind also occurred when teachers used procedural cues to orient students to using external materials to support learning, or toward cognitive operations for storing content. The basic feature characterizing these incongruities between teachers' intentions and students' perceptions of appropriate cognitive responses was that students misperceived the immediacy of operating on content or the nature

of the cognitive product to be created. In these events, students focused on processing the content in the immediate task rather than orienting, as the teacher intended, toward affective states or content to be presented later.

A major feature of students' cognitive responses to instruction was that the clarity with which they perceived the teachers' intentions in instructional stimuli often was inversely proportional to the amount of information which students had to cognitively process. When the focus was a macroscopic unit of content, such as a rule or a complicated cognitive operation, there was considerable variability in the accuracy of students' perceptions of the instructional stimulus. This is yet another affirmation of the principle that people have limited cognitive capacity (see Calfee, 1981). When students were not overly taxed by the newness or complexity of the material, they were relatively more accurate in perceiving the teacher's intentions and reporting that they could successfully carry out the cognitive operations teachers communicated via instructional stimuli.

There is an important qualification to this finding that constitutes a third salient feature of students' cognitive lives in the classrooms sampled for this study. When students had a well-practiced or automatic cognitive response to an instructional stimulus, teachers' signals about how students should cognitively deal with lesson content were clearly perceived. Moreover, students reported they felt very adept at carrying out the teacher's intentions for cognitive processing in these instances. When instructional events were of this type, then students seemed to have little difficulty in wrestling with very complicated cognitive processing such as many-stepped rules or metacognitive monitoring.

A fourth major finding was that the ability of students to undertake cognitive processing that teachers signaled by instructional stimuli depended on their mastery of content at the time. When students were directed to engage in almost any subcategory of operating or consolidating, the extent to which they had comprehended prior lesson content was a fairly good

predictor of the accuracy with which they perceived the teacher's instructional stimulus.

This finding appears paradoxical. If students are less able to use a particular cognitive process that develops knowledge because they do not have that knowledge, how do they ever acquire this knowledge in the first place? As pointed out earlier, a particular instructional stimulus can cue different cognitive responses for different students. Though students might not use the particular cognitive process intended by the teacher, they may use another that produces relatively equivalent knowledge. In this sense, there may not always need to be a one-to-one correspondence between teachers' instructional stimuli and students' cognitive responses. Since we did not collect achievement data, we have no ground for claiming the cognitive response intended by the teacher, and only this cognitive response, is the most effective way for students to learn content. This is another reason why future research on teaching must attend more carefully to the cognitive life of classrooms since most current views of a teacher behavior - are rationalized in terms of only one kind of cognitive response by students.

The third and fourth findings just discussed show the important interaction between factors that cognitive psychological researchers label the state of prior knowledge and the automaticity of executing cognitive plans. When students suffer from a lack of content knowledge needed to proceed with an instructional task, their facility at trying out complex processing tasks cannot compensate for this deficiency because the information to be processed is unavailable. In this case, attention to instructional stimuli probably declines because students' cognitive resources are invested in building representations of new content in short-term memory. Hence, instructional stimuli are unlikely to be strong influences on students' learning when students have underdeveloped knowledge needed as a prerequisite for the content being delivered in lessons. In turn, their reports about teachers' intentions for cognitive processing should be vague or off-target. This is what we observed.

When students have necessary levels of prior knowledge, the automaticity with which they can carry out cognitive strategies for learning also might be related to the probability they accurately perceive teachers' intentions for their cognitions as manifested in instructional stimuli. Automatic cognitive strategies require fewer cognitive resources to execute, especially as regards space in working memory. This unused space may be devoted to any number of tasks, one of which may be making maximum use of guidance provided by instructional stimuli. Thus, students who have prior knowledge that can support new learning and who have automatic cognitive strategies to undertake tasks that promote learning should be in the best position to profit from instructional stimuli. This, too, is what we observed students reporting about instruction.

Also, when something in the instructional environment disrupts automatically executed cognitive strategies, such as an unexpected instructional stimulus that does not fit well into the cognitive strategy, students probably will report confusion about the meaning of the instructional stimulus. This hypothesis was also supported in our data.

The presence of automaticity in learners' cognitive responses leads to a methodological problem. Automatic productions are generally not available to consciousness, by which we mean that people are not very aware of how they are manipulating information. We usually cannot think about the automatic procedure when we perform it without disrupting the procedure to some degree. Thus, stimulated recall methodology as used in this study may have serious limitations for investigating automatic cognitive procedures students use in classrooms.

Indeed, such procedures may introduce a paradox. In some contexts efficient learners, those who employ automatic production, may not be able to give veridical self-reports of their cognitive processes, while less efficient learners may have more access to these procedures simply because they are not automatic. Thus, the findings reported here might be biased in the direction of less efficient learners. The more graphic

examples of cognitive processing on the part of students in this study may have come from those students who were less efficient. Of course, without both achievement data and evidence of cognitive processing during acquisition (i.e., during the actual lessons we videotaped), this remains a speculation.

These analyses and hypotheses also relate to another distinction often made in cognitive psychology between declarative and procedural knowledge. The former constitutes the "static" content of memory that is representative of the structure of the subject matter a student has learned. The latter refers to the cognitive operations underlying the ability to perform an intellectual task. The two kinds of knowledge are not independent. The state of a learner's declarative knowledge will influence the cognitive procedures that he/she will be able to perform (Shavelson, 1981). The relationship between the two forms of knowledge is exemplified in the data reported here by students who were apparently unable to execute particular cognitive processes because they had not learned prerequisite content. But as declarative and procedural knowledge are applied over time, both can change. Thus, the procedural knowledge that is linked to declarative knowledge at a time (X + 1) may not be the same procedural knowledge that served to create that declarative knowledge at time X. In turn, older instructional stimuli that were perceived accurately at time x may lead cognitive processing astray at time X + 1 if they cue inappropriate procedural knowledge.

Finally, students appeared to be strategic in the cognitive processing they used as they attempted to learn from teaching. On those occasions where the teachers did not describe clearly their intentions for students' cognitive processing, and even on some occasions where teachers said they had no clear intentions, students nonetheless made cognitive responses to what they perceived to be instructional stimuli. This is particularly crucial, not only because it relates to a central assumption of the cognitive mediational paradigm, but because it suggests that students will construct meaning for classroom activities

regardless of whether the teacher (or an instructional theorist) does. Failure to incorporate this student variable into research on teaching may lead to misrepresentations of how students learn from teaching, thereby limiting the completeness of theories of teaching. Whether this omission also will limit the ability of researchers to develop useful prescriptive theories of instruction must await experimental evidence.

A significant caveat needs to be made regarding this last conclusion about students' strategically active cognitive processing in classrooms. Stimulated recall methodology does not test whether participants' reports validly reflect the events that are being probed. Indeed, it is possible that participants' responses to questions during stimulated recall are constructions about what they might have been doing or would do if given another chance. These constructions thus may not be the same as actual cognitive processing during the event. More research is needed at this point. Hypotheses about these phenomena should be tested in experiments where students' strategic cognitive processes are operationalized by experimenters to become independent variables (for example, see Winne & Marx, 1980; Winne, 1982 a, b). The findings produced by a methodology like that employed in this study can be used to create hypotheses for such experimental tests, but they cannot confirm or disconfirm these hypotheses.

CONCLUSIONS

This study demonstrates that students mediate instructional events with their cognitive processing, thus supporting the cognitive mediational paradigm as a heuristic for research on teaching. The range of teaching studied in this project was limited because the only lessons we observed were recitation lessons. Obviously, this sample of lessons does not represent the full range of classroom teaching strategies. The cognitive lives of students described here may not generalize to the diverse forms and settings for teaching that can be found in today's schools. Nonetheless, we believe that the phenomena described here provide a useful beginning for investigations of these other settings.

Future research will have to improve on our procedures for soliciting descriptions about cognitive operations from the teachers and students. While occasions where a student or teacher was eloquent and precise in describing cognitive processing provide support for the existence of these phenomena, it is possible that our methodology may have overly structured these descriptions. Although this difficulty is inherent in the procedures for collecting and analyzing data that we used in this study, we hope that the opportunity to improve on the precision of our findings will stimulate further research.

Finally, it is important to emphasize that this first study did not investigate the relations between students' perceptions of instructional stimuli, their competence at executing associated intended cognitive processes, learning, and teaching effectiveness as reflected by students' achievement. This research is extended in the three studies that follow to examine relations between students' perceptions of instructional stimuli and their achievement that results from instruction. This extension is necessary for the cognitive mediational model to provide a basis for prescriptive theories of teaching, that is, statements about the causal relations between teaching events and students' learning that <u>ought</u> to be followed to improve education.

We are convinced from these data, our own previous theoretical and empirical work (Marx & Winne, 1981; Winne & Marx, 1977, 1979), and the work of other researchers in this area (Corno, 1981; Doyle, 1980; Shavelson, 1981; Weinstein, in press) that experiments regarding the cognitive mediational paradigm are warranted. Such experiments must meet a number of requirements in order to test the validity and utility of a cognitive mediational framework for research on teaching.

First, instruction must be designed and delivered in such a way that the teacher's (or researcher's) intentions for students' cognitive processing are clearly communicated to students. We had ample evidence from this study that when teachers hold ambiguous intentions or when they are unable to communicate their intentions clearly, students' thinking can vary substantially.

Second, students need to perceive the occurence of the instructional stimuli that communicate these intended cognitive responses and they have to be able to execute the appropriate cognitive responses when they are needed. This second requirement carries several methodological implications. Instructional stimuli need to be salient. Students should have sufficient practice discriminating the instructional stimuli and executing their associated cognitive response so that these cognitive operations can lead to efficient learning. Evidence must be collected that the cognitive responses are executed as intended. Documentation providing this evidence should be constructed in such a way that, logically, one would have to perform the intended cognitive operation in order to produce the evidence, say notes in a notebook. A final methodological implication emanating from this second requirement is that evidence should be collected of learners' prior knowledge of content to be learned in the study. This relates to the declarative-procedural knowledge relationship discussed earlier.

Third, a program of research should show that the control of students cognitive mediation of instruction can occur in the complex environment of school classrooms, as well as in the more rarefied and controlled environment of the laboratory. This last requirement is the acid test of educational research, and is particularly necessary for research on teaching. One of the major strengths of the process-product model of research on teaching is that most of the empirical work took place in classrooms. While from our vantage point the findings that emerged from the process-product model have been neither theoretically rich nor empirically robust, they have led to recommendations and applications in educational settings (Gage, 1978). If the cognitive mediational model is to supplant the process-product model, it must demonstrate this same utility.

The next three chapters of this report document several experiments concerning students' cognitive mediation of instruction. All took place in schools, although the first two studies are described more accurately as laboratory studies, since

they employed videotape analogs of teaching. The third set of studies was a less controlled but more ecologically valid test of the cognitive mediational model, and took place in five elementary school classrooms over a three-month period.

CHAPTER 3:

Training Students to Use Strategic Responses to Videotape Analogs of Classroom Teaching: Study II

CHAPTER 3

Training Students to Use Strategic Responses to Videotape Analogs of Classroom Teaching

INTRODUCTION

Many models of instruction assume that learners are cognitively active and strategically purposeful as they attempt to learn from teaching. Following on this assumption, one of the teacher's aims in classrooms becomes that of shaping or channeling these cognitive processes in ways that facilitate students' achievement. For example, a teacher who asks a compare-and-contrast question probably intends students to link two concepts according to the attributes they share, and to rehearse the link between attributes of the concepts as they mentally answer the question. It is believed that these cognitive operations help students to store these concepts in memory in a form appropriate for achieving one of the teacher's educational objectives.

Characterizing learning from teaching in this fashion suggests that there are two varieties of stimuli in instruction to which learners can respond cognitively. One variety constitutes the content to be learned, what is traditionally labeled the academic curriculum. The second variety are stimuli that teachers (and other media for presenting curriculum material) use to cue learners to use particular cognitive strategies in order to accomplish learning. These are instructional stimuli. To profit from instructional stimuli presented by a teacher, learners must accomplish three cognitive tasks. First, they must perceive instructional stimuli, i.e., notice their occurrence and understand the cognitive operations or strategies the teacher intends them to engage to facilitate learning. Second, students must carry out the cognitive activities to create or manipulate information that should be stored in memory as a representation of the curriculum to be learned. Finally, they must encode this instructionally prepared content for later retrieval, e.g., on a test.

This analysis of tasks students must pursue to learn from instruction identifies two points where instructional effects may be attenuated. First, if instructional stimuli are not perceived, they can not affect learning. Second, if instructional stimuli are perceived, but the learner does not carry out the cognitive strategies they cue, intended instructional effects are less likely to appear. Only when learners perceive and carry out the intentions behind instructional stimuli, i.e., when the instructional stimulus functions to control the cognitive processing of content as intended by the teacher, can the conditions influencing of students' learning be inferred unambiguously (see Winne & Marx, 1979, 1980). This often uncontrolled variance in learners' actual use of instructional stimuli can account for some weaknesses of validity in current instructional theories (Winne, 1982a).

Several studies (e.g., Bassett & Kibler, 1975; Dansereau et al., 1979; Winne, 1982b) have showed that students trained to accomplish some or all three aforementioned cognitive tasks learned more than untrained students. This study tested whether elementary school students could be trained to perceive and to act on instructional stimuli common in recitation lessons; and whether these cognitive operations would facilitate learning.

The study proceeded in three phases. During the first phase, a random half of the students from two classrooms, each one in a separate school, were given instruction in identifying two different instructional stimuli presented in specially produced videotaped lessons, and in responding differently to each cue with a procedure that integrated cognitive and written components. As will be described more fully later, the written components corresponding to features of intended cognitive responses that students recorded on specially prepared worksheets provided evidence about the degree to which they carried out cognitive features of the complete instructional responses. Then, in phase two, all the students in each of the classrooms viewed one videotaped lesson on the same topic as they had experienced during the training phase, and took multiple-choice and essay tests on

the information in that lesson. In the third phase, all the students in each classroom viewed six lessons about a new topic over a two-week period, and were administered multiple-choice and achievement tests after the sixth lesson and again two weeks later.

METHOD

Participants

Twenty-four 7th-grade students from one class in one school and 25 7th-grade students from one class in a second school were volunteers. Both schools serve middle-class neighborhoods in one of the same school districts as Study I. Within each class, students were randomly assigned to a training group (n = 12 and n = 13) or a comparison group (n = 12 and n = 12) under the constraint that groups were approximately balanced by sex.

Materials

Two curriculum units were created, one on the study of sleep and one about applied psychology (see Appendix A).

The sleep curriculum was used in phases one and two of the study. The psychology curriculum was used in phase three. A unit consisted of 12 segments, each of which presented two concepts. Each concept was defined by two attributes. One defining attribute was common to both concepts in a segment. Each concept also had a second attribute which was unique to that concept. For example, the segment on circadian rhythms in the sleep curriculum defined (1) the concept of bodily rhythms as (a) relating to systems of the body, and (b) having cycles about 24 hours long; and (2) the concept of psychological rhythms as (a) relating to people's behavior and (b) having cycles about 24 hours long.

Accompanying each segment, a text overview gave the title of the segment plus a short descriptive paragraph about the information to be presented. This paragraph introduced the label for each concept presented in the segment. Also, all students received a worksheet for each segment. It presented the title of the segment and for each of the two concepts in the segment: the concept label; the two defining attributes; and two short examples

of the concept, each of which contained representations of the two defining attributes.

Videotapes were pro[illegible] four students were taught about each segment. The [illegible] was [illegible]ripted (as in Clark, Gage, Marx, Peterson, Stayro[illegible] 19[illegible] to be a common recitation lesson involving a te[illegible] overv[illegible]w, t[illegible]acher and student presentation of two [illegible]es p[illegible] concept, teacher questions and elaborative feedback [illegible] students' answers, and a summary (see Appendix B). Th[illegible] instructional stimuli were embedded within the teaching [illegible] segment. Two of these were intended to be functionally [illegible] cues for learners to engage in a set of cognitive and wr[illegible]te[illegible] operations we labeled the consolidating response (described u[illegible] Training). The third instructional stimulus cued students to engage a set of cognitive and behavioral activities labeled the compare-and-contrast response (also described under Training).

Training

Overview. Students assigned to the treatment group were trained to perceive instructional stimuli that signaled them to carry out cognitive and behavioral operations. The instructional stimulus that cued consolidating responses to instruction was a teacher statement or question concerning students' mastery of the defining attributes of a concept such as, "Make sure you understand the key features of bodily circadian rhythms" or "Is that really clear now?" The complete instructional response to this instructional stimulus consisted of both cognitive and written operations: (1) mentally rehearsing each defining attribute two times, (2) reading and analyzing the first example on the worksheet to locate representations of each defining attribute, (3) circling the words in the example that represented each attribute, and (4) drawing a line to connect the circled material with its corresponding defining attribute. Two consolidating responses were cued per segment, one for each of the two concepts.

These same students also were trained to make compare-and-contrast responses to an instructional stimulus that directed them

to consider the shared attribute and the unique attributes of the two concepts in each segment. An example of this instructional stimulus is the question, "Do you see how bodily rhythms and psychological rhythms are the same and different?" The cognitive and behavioral components students were to engage as a response to this instructional stimulus involved a sequence of four tasks: (1) attempt to retrieve from memory the defining attributes of each concept, or find them on the worksheet if this fails; (2) mentally identify which attribute is shared by the two concepts and which are unique to them; (3) place an "o" on the worksheet beside the attribute shared by both concepts and an "x" beside the attributes unique to each concept; and (4) mentally create an answer to the compare-and-contrast question implied or stated in the instructional stimulus.

Training group. A random half of students in each classroom were trained to identify instructional stimuli during the videotape presentation of a segment and to carry out the cognitive and behavioral operations constituting instructional responses to these cues. Training occurred in five 32-41 minute sessions (M = 37 minutes) distributed evenly over two weeks. The first ten segments of the 12-segment sleep curriculum were used for training. The general procedure for training, which took place in one group per classroom, was as follows.

First, as the videotape was played, the trainer identified the first occurence of an instructional stimulus. The tape was stopped and the components of the instructional response were described and demonstrated. Following this, a test item was presented and the trainer described how the instructional response consisted of applying the same cognitive operations to the similar content in order to answer the test item. Consolidating responses were related to multiple-choice items that required students to analyze an example of a concept as they had to do on the worksheet. Compare-and-contrast responses were related to essay items asking students to produce an answer describing similarities and differences between the two concepts in a segment.

Next, students watched the videotape and were cued by the

trainer when the instructional stimulus was delivered. The students then attempted the instructional response. The trainer provided corrective feedback on this attempt after students answered a test question about the content just manipulated according to the cognitive-plus-behavioral response strategy. Then the test item was reviewed and the trainer reiterated how it could be answered by using the cognitive components of the instructional response associated with the instructional stimulus.

Similar practice and feedback continued for the remaining segments. Consolidating responses were introduced on the first day of training and practiced thereafter. Compare-and-contrast responses were introduced on the third day of training and practiced on that day and the fourth day. The following sections provide a detailed description of each training session delivered to the two training groups by a single trainer. Appendix C presents the scripts used by the trainer to deliver each training session.

The objectives for the first training session were to: (1) train students to discriminate the instructional stimulus for consolidate instructional responses (e.g., "Make sure you understand about hypersomnia."); (2) introduce and train students to carry out the intended cognitive components for consolidating information; (3) train students to provide observable written indicators corresponding to their cognitive activities; and (4) motivate students to carry out both the cognitive and behavioral operations by showing them the link between these activities and multiple-choice items that they would take on later achievement tests.

The session began with a general introduction to the experiment and a description of the procedures that would be followed over the remaining sessions of training. Specifically, students were told that videotaped lessons would be shown to them, and that the trainer would instruct them in ways to think and take notes that would promote learning the information taught in the lessons. They were informed that they would have opportunities to practice a thinking strategy, and that they would answer test

items during training that were based on the information presented in the videotape. They also were told that they would see two short "lessons" (i.e., segements), per day on the videotape, and were strongly encouraged to imagine that they were in the classroom in which the videotape had been made.

Following this introduction, the students read the overview of the first segment on sleep. At this time, they also were reminded that when they finished the reading, they should imagine that they were in the classroom displayed on the videotape they would watch next. At a point in the videotape when the teacher said, "So is it clear what I mean by bodily rhythms?", the trainer stopped the videotape and pointed out the instructional stimulus to students. He described that it was a signal the teacher was giving to tell students to think in a special way. The trainer then referred the students to the worksheet corresponding to the concept referenced by the teacher's instructional stimulus. The cognitive components of the consolidate response were then described to them, and the trainer modelled for the students how and what they should be thinking by talking aloud through the procedure. The trainer then described the behavioral indicator of this cognitive strategy by referring students to another worksheet on which one half of the written response had been drawn in for students (i.e., circling the words in the example that represented one of the concept's defining attritutes, and connecting that circle with the attribute printed on the worksheet; see Appendix C). Following this first introduction, the trainer reviewed the entire sequence of instructional stimulus, intended cognitive response, and behavioral indicator.

For the next worksheet, corresponding to the second concept, this same procedure was followed except that students were required to answer a multiple-choice test item immediately after the videotape had been stopped and before the trainer discussed the intended cognitive response and the written indicators of cognitive operations to be recorded on the worksheet. As the worksheet was discussed, the trainer pointed out the parallels between the cognitive operations students were to carry out during

teaching and the thinking they needed to do to answer the multiple-choice question they had just taken. Following this description, the trainer told students to turn to the multiple-choice item and described for them exactly how the intended cognitive response helped them answer this kind of question. These descriptions were intended to provide an incentive for students to use the cognitive response during teaching. The answer to the multiple-choice question was then provided to students.

In the second lesson (corresponding to the second segment of the sleep curriculum) presented on this first day of training, students were given an opportunity to practice the intended cognitive response and written indicators. This was followed by feedback from the trainer concerning the correctness of what students had circled in the examples they analyzed on the worksheet, and further description about how using the intended cognitive response would help students answer the test item correctly.

The second training session began with an overview of the consolidate intended cognitive response and behavioral indicators that had been trained in the first session. Students were reminded that doing these activities would help them answer multiple-choice test items like those they had experienced in the preceding session. They were further reminded to try to imagine that they were "right in the classroom" they were seeing on the videotape. As well, the trainer reminded them specifically of the kinds of instructional stimuli the teacher on the videotape used to signal them to engage in the cognitive and behavioral operations described earlier.

Following this introduction, the training session followed essentially the same routine as the preceding one except that the videotape was not interrupted when the instructional stimulus was delivered. This required the students to respond within a shorter time span, as they would have to do in all subsequent training sessions and in later phases of the study. In this session, none of the worksheets that students used had been marked previously to

illustrate the behavioral indicators for the consolidate intended cognitive response.

The trainer began the third training session by reviewing the entire sequence of instructional stimulus, intended cognitive response and behavioral indicator, plus the relevance of the intended cognitive response for correctly answering multiple-choice quiz items.

Following this introduction, the videotape was played, and students responded to the consolidate instructional stimulus as in session two. When the instructional stimulus for making a compare-and-contrast response appeared (e.g., "Do you see the difference between dreams you tend to remember and dreams you tend to forget?"), the trainer stopped the tape. He described what the instructional stimulus was and illustrated the intended cognitive response by talking through the cognitive operations the teacher intended students to use when he provided this instructional stimulus. Students were next shown the behavioral indicator for this intended cognitive response. Then the trainer had students turn to the next page of their booklet on which an ideal answer had been written for an essay question that asked students to compare and contrast these two concepts. The trainer explained how this second intended cognitive response would be helpful in answering short essay questions that required students to compare and contrast two of the concepts presented in the lesson. After this introduction, students responded to the multiple-choice item that was associated with content referred to by the preceding consolidate instructional stimulus. Then, they received feedback about their behavioral indicators connected with this instructional response and about their answers to the multiple-choice test item.

Students were then shown the next part of the videotape and told to practice looking for both types of instructional stimuli. The videotape was not interrupted for the instructional stimulus, intended cognitive response, and behavioral indicators associated with the consolidate strategy. It was stopped momentarily (approximately 30 seconds) following delivery of the

compare-and-contrast instructional stimulus to allow students time to carry out the newer intended cognitive response and behavioral components this stimulus signaled. Following this part of the videotape, students responded to the multiple-choice item drawn from this second part of the videotape. Another compare-and-contrast essay question and its ideal answer were analyzed in terms of how the compare-and-contrast intended cognitive response would help students answer these kinds of questions. Then both types of behavioral indicators students had marked on their worksheets for the second part of the videotape were reviewed while the trainer provided feedback. Finally the answers to the multiple-choice question were reviewed as the trainer reiterated how making the consolidate intended cognitive response should help students answer this type of question.

The objectives for the fourth training session were to have students continue to practice identifying both types of instructional stimuli, making both types of intended cognitive responses, and providing the behavioral indicators associated with each intended cognitive response. This session began with the trainer reminding students to imagine they were actually in the classroom they were seeing. Then, students read the overview, watched the videotape, responded to the essay and multiple-choice test questions, (in that order to avoid providing information for the essay answer in the multiple choice item) received feedback on their behavioral indicators, and then received feedback about their answers to the test questions. During this last activity, the trainer emphasized how making the intended cognitive responses during the lesson would enhance students' performance on both types of test items.

In the fifth and last session of training, students were given another opportunity to practice all of the skills they had been taught in the preceding sessions. Students read the overview, watched the videotape, answered the test questions, received feedback on their behavioral indicators for the two types of intended cognitive responses, and analyzed their answers to the test questions in terms of their intended cognitive responses.

The trainer again stressed that making the intended cognitive responses should help them get better scores on achievement tests, and that students had to pav close attention to the teaching displayed on the videotape.

Comparison group. A second trainer provided the remaining half of students in each class with exactly the same materials (videotapes, worksheets, and test items). Students in these groups did not receive training to identify instructional stimuli in the videotaped lessons or to make instructional responses and the associated behavioral indicators for these cognitive operations. Otherwise, all other events were the same, including the amount of time they spent viewing the videotapes and discussing the topics (range = 31 to 41 minutes, M = 35.5 minutes).

Instruments

Aptitude. Two weeks before training, all students were administered a standardized vocabulary test (Cognitive Abilities Test, Thorndike & Hagen, 1971) to gauge verbal ability.

Use of intended cognitive response. Two measures of the effects of training were obtained two days after the last training session (session 5). Both were based on the remaining two segments of the sleep curriculum. A consolidate score and a compare-and-contrast score were obtained by examining the way trained students marked their worksheets as they had been taught. These scores were assumed to reflect students' use of the intended cognitive response components of the complete instructional response. The basis for this assumption is as follows.

For the consolidate instructional response, we reasoned that a student must analyze the way the attributes of a concept are represented in the example provided on the worksheet in order to circle the words that represent those attributes. Then, in order to draw a line that connects this circled material to the statement of an attribute, the student must read the attribute to verify that a proper connection is being drawn. This act of reading is equivalent to rehearsing the attribute.

The links between the cognitive and the written components of the compare-and-contrast instructional response are more tenuous. For the student to mark the concepts' attributes correctly, the attributes must be read. As before, this constitutes at least one rehearsal of the attributes. As well, to identify the two concepts' shared attributes, a mental judgment of similarity must be made. The remaining two cognitive components of the compare-and-contrast intended cognitive response are not clearly reflected in the behavioral indicators of this instructional response. Students who followed our training would have made a mental judgment of dissimilarity to mark the two concepts' unique attributes. However, an alternative to this intended cognitive activity is to mark the remaining attributes without making this judgment or even reading them, since they must be the unique attributes by elimination. There is no behavioral indication in our data that students complied with the remaining cognitive component, namely, mentally framing an answer to a compare-and-contrast question. We can only assume that students were motivated to carry out this activity because they were shown how it would transfer to the short essay test items.

The same kinds of scores were obtained from the worksheets students had available for the psychology curriculum which was delivered over four consecutive days (Monday-Thursday of the following week). The 12 curriculum segments were presented on videotape, partitioned into three segments per day.

Achievement. An eight-item multiple-choice test and a two-item essay test provided scores describing students' achievement on these last two segments of the sleep curriculum (see Appendix D). A 24-item multiple-choice test and a four-item essay test (Appendix D) based on the psychology curriculum were administered on Friday, the day after all of the students had viewed the last three segments of the psychology curriculum. The multiple choice and essay tests for the psychology curriculum were readministered to all students two weeks later.

For both the sleep and psychology achievement tests, essay items were given first in a separate booklet and collected before

the multiple-choice items were given out. As each student finished the essay items, he/she raised his/her hand, at which time the tester collected the essay exam booklets and distributed the multiple-choice test booklet.

Scoring procedures. For the consolidate score, one point was awarded if the student encircled the part of the example that represented a concept's attribute and drew a line connecting the circle to the attribute listed on the worksheet. Both the circle and the line had to be present and unambiguous for credit to be awarded. Two points were possible for each concept (one for each of the two attributes), four points for each segment (two concepts with two attributes each). For the sleep posttest, the maximum consolidate score was eight because there were two segments for which behavioral indicator data were available. For the psychology curriculum, the maximum consolidate score was 48 because there were 12 segments for which behavioral indicator data were collected.

The compare-and-contrast behavioral indicator score was derived by awarding one point for indicating which of the attributes for the two concepts in a segment were the same, and one point for indicating the two attributes that were unique. In the latter case, the unique attributes for each of the two concepts had to be clearly indicated. Thus, there were two possible points for each segment, one each for indicating the shared and the unique attributes. For the sleep curriculum, the maximum score was four (2 segments x 2 points); for the psychology curriculum the maximum score was 24 (12 segments x 2 points).

Multiple-choice items were scored as either correct or incorrect. The number of items correct served as the multiple-choice test score.

Seven different scores were recorded for each essay item. The first two scores for an item each had a possible value of zero or one. These two scores indicated whether the unique attribute of each of the two concepts had been retrieved correctly. The sum of these two scores for each item aggregated over the two essay items was calculated to produce the attributes recalled scale. Scores on this scale thus ranged from zero to eight.

The third score indicated the degree to which the similarity between the two concepts had been specified correctly. This score was zero, one, or two. One was given if the idea of a similarity was mentioned, and two was given for correct specification of the similarity, i.e., describing the shared attribute. Zero indicated no recall or completely incorrect recall.

The fourth score ranged in value from zero to three and referred to the correct specification of the difference between the two concepts. A score of one was given if the idea of a difference was mentioned. Another point was given for correctly linking each unique attribute to its concept label when identifying the difference between the two concepts. The sum of these third and fourth scores for each item aggregated over the two essay items was calculated to produce the compare-and-contrast scale. Scores on this scale could range from zero through twelve.

The fifth score, having a value of zero or one, awarded a point for including an example or elaborative incidental content.

The sixth and seventh scores reflected two types of errors. One type of error was an interchange of attributes across function, that is, referring to a similarity as a difference. The second type of error was an interchange of attributes across concepts, that is, defining one concept from a segment using the unique attribute of the second concept in that segment, or either attribute of a concept in another segment. Each of these scores had possible values of zero or minus one.

The sum of all these scores produced a total score for each essay that could range from zero to eight. For the sleep curriculum, aggregating all seven of the scores on the two-item test produced a total essay scale with a range from zero to 16. Total essay scores could range from zero to 32 for the four-item psychology essay test. (A detailed scoring manual for the essay tests is provided in Appendix E.)

RESULTS AND DISCUSSION

For all following analyses, the original sample is reduced to 37 (16 total in the trained group, 21 total in the comparison

group). The joint criteria for inclusion in the analyses was that students had to be present for the achievement tests on the sleep curriculum, the four days when the psychology curriculum was presented, the immediate posttests on the psychology curriculum, and the delayed psychology achievement tests.

Caveats

It is essential to stress that several aspects of interpretations of the multiple regression analyses reported below must be considered very tenuous. First, it is reasonable to raise the question of what population is supposedly represented by the samples we were able to work with in this project. Second, the representativeness of whatever population these samples describe is limited by the small sample sizes we used in order to carry out the very labor-intensive treatments. Third, the stability of all statistics based on these small samples is doubtful. Although we follow conventions about reporting probabilities of type I errors common in research on teaching, we strongly advise against placing too much faith in "statistical significance." Overall, we encourage that these analyses be taken as descriptive of the groups we worked with and the processes we explored rather than as a strongly probabilistic basis for inferences about effects in a population.

The third problem, i.e., instability, warrants special mention whenever aptitude-treatment interactions are reported because of the extremely small numbers of cases on which separate regression slopes depend. In many of these instances, especially in the analyses described later for Study IV, such interactions probably reflect the influence of only a handful of students. We have reported these analyses, however, because the effects that were observed merit consideration in future research. Nonetheless, we do not claim that these results are robust. Replication studies need to be undertaken to explore the results we present, and thereby test the speculations we offer here.

Can Students be Trained to Make Intended Cognitive Responses?

The worksheets students had available during the instructional sessions were scored to reflect their use of the

consolidate and the compare-and-contrast responses. There would be no reason to expect students in the comparison group to exhibit written indicators on their worksheets of having made these two kinds of cognitive responses to the instructional stimuli presented in the videotaped teaching. In fact, this was the case for both the sleep and the psychology curricula.

Students in the groups trained to use these intended cognitive responses and to reflect their cognitive operations on their worksheets generally did so (see Table 3). For the sleep curriculum, trained students correctly executed, on average, seven of eight indicators of consolidating, and all students scored perfectly on the compare-and-contrast behavioral indications. Transfer of this strategy to the psychology curriculum for trained students was successful. On average, they used about three out of a possible four behavioral indicators of consolidating per segment, and they were nearly perfect at providing indicators for the compare-and-contrast cognitive operations.

Another measure of transfer of training is the correlation between the behavioral indicator scores for trained students across the curricula. This was .44 ($p = .03$) for the consolidate scores. The seemingly low value is partly attributable to a restricted range of scores from the sleep curriculum. No correlation can be calculated for the compare-and-contrast scores because there was no variance for these scores on the sleep curriculum.

Interestingly, the correlation between consolidate scores and compare-and-contrast scores for the psychology curriculum was -.52 ($p = .01$). We have no hypothesis about why this inverse relation appeared.

Overall, then, these results provide evidence that students can be trained to make cognitive responses to instructional stimuli in accord with a teacher's or researcher's intentions. Moreover, students also can provide observable indications of whether they actually use these cognitive responses during instruction. The short-term transfer of these skills was

TABLE 3

Means, Standard Deviations, and Maximum Scores for Behavioral Indications of Intended Cognitive Responses for Trained Students
Study II

Curriculum	Consolidate	Compare-and-Contrast
Sleep	6.94 (8)	4.00 (4)
	1.51	0.00
Psychology	34.33 (48)	22.78 (24)
	5.20	1.70

Note: Upper numbers are means, lower numbers are standard deviations. Numbers in parentheses are maximum scores.

demonstrated here, but we believe that long-term transfer almost surely would rquire some kind of sustained practice coupled with an environment that provided incentives for students to maintain these responses to teaching. This need is implied by the decrease in trained students' use of these skills during the psychology curriculum, particularly with regard to the consolidate instructional response.

How Do Intended Cognitive Responses Relate to Achievement?

To gauge the relation between students' use of intended cognitive responses and achievement, correlations were computed between behavioral indicator scores and achievement scores for trained students (see Table 4). No statistically reliable relationships were observed among the psychology behavioral indicator scores and psychology achievement scores ($p < .10$). (Recall that correlations involving the sleep consolidate scores or either psychology intended cognitive response scores are attenuated due to restricted ranges.) In sharp contrast and despite the possibility of attenuated variance, the behavioral indicator scores for consolidate responses in the sleep curriculum show moderate positive and fairly consistent relationships with all measures of achievement.

This finding is baffling. Among trained students, those who displayed more behavioral indicators (and presumably employed the intended cognitive response more frequently) while viewing the two videotaped lessons upon which the sleep achievement tests were based, tended to have higher scores on all three essay scales and one of the three multiple-choice scales. Yet this relationship did not hold for behavioral indicators completed during the psychology curriculum. Recall that the correlation between consolidate scores across the curricula was only .44. While this low to moderate correlation may be an artifact of attenuated variance, it is also possible that students' use of intended cognitive responses did not transfer optimally from the sleep curriculum to the psychology curriculum. It may have been the case that the more capable learners acquired the intended

TABLE 4

Correlations among Behavioral Indicator Scores and Achievement for Trained Students Study II

	Behavioral Indicators			
	Sleep		Psychology	
Achievement Measures	Consolidate	Compare-and-contrast	Consolidate	Compare-and-contrast
<u>Sleep</u>				
Multiple-choice	24	-	-	-
Attributes recalled	43*	-	-	-
Compare-and-contrast	50*	-	-	-
Total essay	36	-	-	-
<u>Psychology</u>				
Multiple-choice	57*	-	35	08
Attributes recalled	52*	-	30	00
Compare-and-contrast	57*	-	13	16
Total essay	59*	-	20	10
<u>Delayed Psychology</u>				
Multiple-choice	20	-	-14	16
Attributes recalled	60*	-	22	21
Compare-and-contrast	57*	-	10	33
Total essay	60*	-	12	25

Note: No correlations can be calculated involving Sleep compare-and-contrast indicator scores because of zero variance on this measure. Decimals have been omitted on correlations.

* $p < .05$, two-tailed

cognitive response during training and then failed to use it, or perhaps failed to record behavioral indicators of their cognitive responses to instructional stimuli during the psychology curriculum. However, without data regarding individual differences in aptitude that describes attention to instructional stimuli and perception of intended cognitive responses, this interpretation remains speculative.

Do Intended Cognitive Responses Enhance Learning from Instruction?

To test the relative influence on achievement of training students to make intended cognitive responses, comparisons were made to the group of untrained students.

A backward selection multiple-regression procedure[1] was used to make this comparison. The backward selection procedure begins by entering all possible predictors in the regression equation as one set, and then removes predictors individually if they fail to meet a predefined criterion. The criterion used here was that residualized predictors were removed from the multiple regression equation if the F-statistic testing the proportion of variance accounted for by a predictor had a type I error exceeding .10. Thus, this regression analysis retains only residualized predictors that account for a statistically reliable (at p .10) portion of variance in the dependent variable. Moreover, this procedure has the property that, following the elimination of all statistically unreliable predictors, the remaining predictors are mutually residualized for one another, thereby making them

[1] Throughout this report, we will refer to predictors that remain in regression equations by the name given them in their respective sections that describe the methods for each experiment. We adopt this practice for simplicity of expression. It must be remembered that the correspondence is not perfect between the operational definitions of variables that yield raw data and the statistically defined residual variables contained in a regression equation (see Winne, 1983c). In other words, the construct which we imply by using a label for a variable's operational definition is not isomorphic to the construct reflected by a variable's label after the raw data are residualized in the process of doing a multiple regression analysis. Resources and practical constraints prohibited us from following procedures Winne recommended that might have lessened the potential difficulties entailed by this lack of correspondence.

orthogonal. To express the extent to which each residualized predictor accounts for variance in the dependent variable, we report squared part (or semipartial) correlations. These are correlations between the predictor, as residualized in the context of other predictors remaining in the regression equation at the point where no other predictors can be removed according to the statistical criterion ($p. < 10$), with raw scores on the dependent variable.

Three major categories of predictors were entered at the first step in the backward selection regression analyses. One was a measure of the students' verbal ability, their vocabulary score deviated about the grand mean. The second category of predictors consisted of two effects-coded vectors. One compared the trained and comparison groups, and the other contrasted the two classrooms from the two schools in this study. For the comparison of experimental groups, the group that received training was coded +1 and the comparison group was coded -1. The contrasts between schools cannot be given any particular meaning because we do not have sufficient information regarding the qualities that differentiate them. However, such variance can represent educationally important effects on learning. Also, leaving variance due to schools in the residual mean square would reduce statistical power and confound reliable variance with error variance. Thus, the contrast between schools was included to reduce the residual mean square in the regression analyses, thereby increasing statistical power for the tests contrasting the training and comparison groups. A third category of predictors were aptitude-group interaction terms created by multiplying the vector representing vocabulary and the effects-coded vectors.

Twelve separate analyses were done, one corresponding to each of four types of achievement (multiple-choice, attributes recalled in the essay items, compare-and-contrast statements in the essay items, and total essay score) for the sleep curriculum, the immediate posttests on the psychology curriculum, and the delayed psychology tests.

Descriptive statistics are displayed in Table 5. The results of the regression analyses and effect sizes for differences between groups (difference between adjusted means divided by the square root of the residual mean square from the regression analysis) are presented in Table 6. Group differences are reported if they describe effects that would be statistically reliable at least ten times per hundred comparisons ($p < .10$). More than traditional allowance for type I error was adopted here for exploratory purposes and to increase statistical power.

Training students to make cognitive responses we intended them to make during instruction had some benefits, particularly on the essay tests for the sleep curriculum, but there clearly were no large and consistently positive effects. The effect size statistics indicate that an average student from the comparison group (at the 50th percentile) who received training would have increased scores, relative to the comparison group distribution, on the sleep attributes recalled scale to the 68th percentile, on the sleep compare and contrast scale to the 70th percentile; and on the psychology multiple-choice test to the 60th percentile. No reliable aptitude-treatment interactions were observed. While there are several statistically reliable differences between schools and aptitude-school interactions, these do not lend themselves to any useful interpretation as explained earlier.

It is reasonable to question whether students who participated in training to make intended cognitive responses but who failed to manifest these behaviorally while studying the psychology curriculum should be distinguished from those who participated <u>and</u> evidenced consistent use of these cognitions (as indicated by the behavioral indicator scores). The general absence of statistically reliable relations between behavioral indicator scores and achievement scores (see Table 4) suggests that selecting students for their performance levels on behavioral indicators of intended cognitive responses would not alter the direction of the preceding findings. In addition, the behavioral indicator scores on the sleep curriculum clearly show that all the students had mastered the cognitive operations they had been taught.

TABLE 5

Means, Standard Deviations, and Correlations Among Vocabulary and Achievement Variables
Study II

	M	sd	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Vocabulary	21.63[1] 22.62	4.62 4.78	76[2]	50	31	39	37	62	52	45	52	44	49	49	51
2. Sleep Multiple-choice	5.50 5.76	2.10 1.87		71	51	62	63	57	48	50	56	49	44	55	53
3. Sleep Attributes Recalled	4.13 2.95	2.16 2.56			84	87	95	65	45	44	50	63	52	50	53
4. Sleep Compare-and-Contrast	6.06 4.67	2.62 2.50				67	93	62	56	61	65	62	59	64	63
5. Sleep Total Essay	10.25 8.29	4.64 4.78					78	68	53	54	60	64	59	60	62
6. Psychology Multiple-choice	17.50 16.43	2.58 4.11						85	71	55	69	87	70	63	71
7. Psychology Attributes Recalled	5.13 4.14	3.12 3.68							80	87	93	72	84	81	85
8. Psychology Compare-and-Contrast	8.13 4.29	6.71 4.23								78	96	61	72	85	80
9. Psychology Total Essay	14.44 7.25	11.81 8.00									85	72	78	87	85
10. Delayed Multiple-choice	17.50 16.48	2.61 4.01										88	70	68	73
11. Delayed Attributes Recalled	4.56 4.10	3.42 3.73											85	87	95
12. Delayed Compare-and-Contrast	7.06 7.24	4.43 4.16												80	97
13. Delayed Total Essay	12.06 12.19	7.60 7.87													87

[1] Upper numbers describe the trained group (N=16): lower numbers describe the comparison group (N=21). Means are unadjusted for vocabulary.

[2] Alpha reliability coefficients are displayed in the diagonal. Decimals are omitted in these and the correlations.

TABLE 6

Backward Selection Regression Analyses
Study II

Dependent Variable	Predictors	[illegible]	F's[1]	p	Variance[2]	F
Sleep						
Multiple choice	Total equation		11.62	< .01	.23	–
	Vocabulary	.50	3.41	< .01	.50	
	Constant	(5.65)	20.09	< .01		
Attributes Recalled	Total equation		3.54	.04	.12	.46
	Vocabulary	.34	2.16	.04	.11	
	Group	.28	1.77	.09	.08	
	Constant	(3.55)	9.39	< .01		
Compare-and-contrast	Total equation		5.70	< .01	.21	.52
	Vocabulary	.43	2.85	< .01	.18	–
	Group	.31	2.10	.04	.10	
	Constant	(5.78)	13.95	< .01		
Total essay	Total equation		5.51	.02	.11	–
	Vocabulary	.37	2.35	.02	.14	
	Constant	(9.14)	12.39	< .01		
Psychology						
Multiple choice	Total equation		12.76	< .01	.40	
	Vocabulary	.64	4.91	< .01	.41	
	Group	.22	1.70	.10	.05	.26
	Constant	(17.00)	37.35	< .01		
Attribute recalled	Total equation	.33	7.41	.01	.34	
	Vocabulary	-.28	-2.15	.04	.08	
	School	-.30	-2.03	.05	.07	–[3]
	Vocab x School	(4.19)	2.00	.05	.07	
Compare-and-contrast	Total equation		9.88	< .01	.33	
	School	-.32	-2.34	.03	.10	
	Vocab x School	-.53	-3.86	< .01	.28	–
	Constant	(6.65)	11.19	< .01		
Total essay	Total equation		10.16	< .01	.43	
	Vocabulary	.27	1.91	.07	.06	
	School	-.33	-2.59	< .01	.11	
	Vocab x School	-.41	-2.91	< .01	.13	–
	Constant	(11.87)	11.90	< .01		

Dependent Variable	Predictors	β	F/t [1]	p	% variance [2]	E
Psychology Delayed						
Multiple choice	Total equation		8.70	<.01	.30	
	Vocabulary	.36	2.52	.02	.12	
	School	-.40	-2.78	<.01	.15	-
	Constant	(16.66)	34.27	<.01		
Attributes recalled	Total equation		11.33	<.01	.22	
	Vocabulary	.49	3.37	<.01	.25	
	Constant	(4.30)	8.33	<.01		
Compare-and-contrast	Total equation		7.52	<.01	.27	
	Vocabulary	.37	2.30	.03	.11	
	Vocab x School	-.28	-1.79	.08	.06	
	Constant	(6.94)	11.42	<.01		
Essay Test	Total equation		12.35	<.01	.24	
	Vocabulary	.51	3.52	<.01	.26	
	Constant	12.14	11.07	<.01		

[1] F-statistics are reported for the regression equation containing all predictors remaining in the equation (p < .10); t-statistics are reported for each predictor's slope coefficient.

[2] The figure reported for the total equation is R^2 adjusted for shrinkage. Statistics associated with each predictor are squared part (or semipartial) correlations.

[3] Effect sizes are not reported for predictors involving differences between schools because we have no way to meaningfully interpret these differences.

A number of possible interpretations of these results can be offered. One interpretation focuses on the effects of the perceptual components of the training, assuming there was little impact on the cognitive strategy component. Training students to perceive the intended cognitive responses associated with instructional stimuli has little transferable impact on students' learning. Thus, it could be claimed that the perceptual training to attend to instructional stimuli increased the salience of content taught during the sleep curriculum, setting the stage for trained students to use whatever cognitive strategies they had developed prior to training on content that was easier to identify as important. As a result, the students developed more elaborate networks in long term memory for the sleep content. This aided them in the relatively free recall retrieval task of the essay test. But for the recognition task on the multiple choice test, the more elaborate semantic networks of the trained students did not provide them with an advantage vis a vis the comparison group. This difference in performance in free recall versus recognition tasks has been documented numerous times in research on long term memory (Anderson, 1980).

A similar explanation of the results can be given focusing on the cognitive strategy training. It is possible that the perceptual training was inconsequential, but that the cognitive strategy training was successful. Thus, in learning the content taught in the final two sleep segments, the trained students employed the cognitive strategies that we had taught them, leading to more elaborate semantic networks. However, if perceptual training was inconsequential, the cognitive strategy may have been used at times throughout the lessons that were not cued specifically by the instructional stimuli. From here the explanation parallels the previous one, namely, that trained students developed relatively more elaborate networks of information but did so under their own guidance rather than taking cues from instructional stimuli.

Neither of these explanations adequately deals with the lack of transfer of scores for the behavioral indicators observed

during presentation of the psychology curriculum. Again, however, the lack of correspondence between the frequency of use of the cognitive strategy and achievement in psychology is puzzling. One explanation is that the use of the cognitive strategy with the psychology curriculum was not helpful. But if so, why were there positive and statistically reliable correlations between behavioral indicator scores and achievement on the tests for the sleep curriculum when the structure of these two curricula was the same? It may be that different students used the behavioral indicators more frequently on the two curricula as suggested by the correlation between the two behavioral indicator scores ($r=.44$). Some individual differences construct other than verbal ability may account for these results. We have found in another study (Marx, Winne, & Howard, 1982), however, that at least the cognitive style constructs of field dependence-independence and locus-of-control appear to be unrelated to children's ability to perceive and execute a teacher's intentions for students' cognitive processing. These two constructs, then, could be eliminated as candidates.

CONCLUSIONS

Further consideration of these results will be taken up following the results of the next study which paralleled many features of this one.

CHAPTER 4

Further Questions about Students' Responses to Teaching in Varying Analog Settings: Study III

CHAPTER 4

Further Questions about Students' Responses to Teaching in Analog Settings

INTRODUCTION

The results of Study II indicate that elementary school students can be taught to use cognitive strategies for classroom learning. This interpretation rests on the assumption that the behavioral indicators represent valid traces of the cognitive operations students were taught to apply to information provided on worksheets. For the purposes of Study III, as in Study II, we choose to interpret the data about behavioral indicators in this manner. If one accepts this assumption, what remains to be explored is the impact of using this cognitive strategy on subsequent achievement.

In an earlier study (Winne & Marx, 1980), we found that university students could easily learn a multi-component cognitive strategy to use in response to instructional stimuli delivered in lectures. But, when students returned to an ecologically valid learning environment (the lecture hall) where the outcome of their learning was important to them (namely, their final grades in an undergraduate course), many students chose not to use the strategy they had been trained to use, but to revert to their own idiosyncratic cognitive responses to teaching that they had used before we trained them in ours. In addition, use of the strategy after only one hour of training had a detrimental effect on learning compared with control groups.

In that study, we speculated that, for many students who tried to use the strategy we had trained, working memory capacity was strained by the requirement that they retrieve, sequence, and execute the components of a new cognitive strategy while they were simultaneously attempting to attend to information being delivered at a rapid pace in the lecture. What these students told us in conversations following formal data collection was that, although

they saw much merit in the cognitive strategies we had trained them to use, they were not proficient enough at carrying out those strategies to record in their notes all the content that was being presented in the lecture.

There probably were two interlocking reasons for this difficulty: The first is that we did not provide the students in our 1980 study with nearly enough opportunity to practice the new cognitive strategies they were to engage following the delivery of instructional stimuli during lectures. This flaw in our procedures resulted in low levels or perhaps even no automaticity in the students' ability to execute the intended cognitive responses which they had learned in training. Thus, the cognitive strategies we trained quite likely were not applied to concepts as thoroughly and expertly as we intended. In turn, the hypothesized gains in achievement relative to a control group would fail to appear under these conditions because the strategy was not used properly.

The second reason that our 1980 study may have failed to yield the findings we expected may relate to students' ability to pay attention to the lecture. Attention is a limited cognitive resource (e.g., see Calfee, 1981; Anderson, 1980) that can be channeled to external information, such as concepts being presented in a lecture, and to internal cognitive activities, such as retrieving and sequencing procedures for learning concepts being presented in a lecture. When attention is focused primarily on external information, there will be little of this resource that remains available to monitor internal cognitive activities if the information being presented externally is complicated, unfamiliar, or fast paced. Alternatively, if internal cognitive activities are not automatic, and therefore demand considerable attention to carry out, not much of one's attentional resources will be available to devote to perceiving information emanating from external sources. The external information that is not attended to could be curricular, instructional stimuli or both. Under these conditions of low automaticity, then, the poor showing of trained students on achievement tests are potentially

attributable to both less curricular information being perceived and less instructionally appropriate cognitive processing of the curricular information that is perceived due to students' inattention to instructional stimuli.

In designing the training procedures for Study II, we sought to insure that students would achieve automaticity in executing the consolidating and the compare-and-contrast intended cognitive responses. The behavioral indicator scores for the immediate test of training on the last two segments of the sleep curriculum showed that we were reasonably successful in this respect, perhaps attributable to any or all of: (1) sufficient length of training (5 sessions), (2) practice and feedback about executing the intended cognitive responses, (3) incentive for learning the intended cognitive responses because of their close association with cognitive activities involved in answering test questions, and (4) training using a format and medium that matched that which students would experience in the tests of how training influenced achievement (i.e., worksheets and videotaped teaching). Thus, Study II demonstrated that students could be trained to use a new cognitive strategy when cued by instructional stimuli. These talents, however, were not as robust as we had expected with respect to their transfer to a longer curriculum, nor were they as important in influencing achievement scores as hypothesized.

A long history of research supports the necessity of the first and second features, namely, sufficient exposure including practice plus feedback, in training programs designed to promote expertise in tasks somewhat like those we taught in Study II (e.g., see Brown, 1978). We judged on logical grounds that the third feature of the training program, relating to providing an incentive for using the cognitive response strategies, could not be a factor limiting the robustness and relevance to achievement of intended cognitive responses like those we trained. Thus, by elimination, we sought to explore the fourth feature of the training environment as a source for the absence of effects predicted in Study II. Our strategy was to decompose the videotape and worksheet components of the training to test the

videotape and worksheet components of the training to test the hypothesis that they might interact with one another in a manner that somehow interfered with transfer of the cognitive strategies or inhibited their positive influence on learning.

This was accomplished by creating a total of three treatment groups and two comparison groups. The three treatment groups all involved training of the cognitive strategies, one with only the videotapes, one with only the worksheets, and one with both the videotapes and worksheets. This last group was the same as the treatment group in Study II. In the two comparison groups the students were exposed to either the videotape only or videotape-plus-worksheets, but they were not taught to use a particular cognitive strategy. The Videotape-plus-worksheet comparison group included the same activities as the comparison group in Study II. A comparison group that only read the worksheets was not included since this would involve only minimal instruction and would have reduced group sizes to extremely low numbers.

METHOD

Participants

A total of 70 seventh-grade students from three classrooms in one school were volunteers in this study. The school serves a lower middle-class neighborhood in the same school district as that which participated in Study II. Students from each of the three classrooms were pooled to form a single population and then randomly assigned to one of three treatment or two comparison groups under the constraint that groups were approximately balanced by sex and for their representation from each of the three original classrooms.

Materials

The same curriculum units, worksheets, videotaped lessons, and tests used in Study II were used here (see Appendices A-E).

Training

Three of the five groups in this study received training to make intended cognitive responses. One of these groups, hereafter referred to as the worksheet-plus-videotape training group,

received exactly the same training as the group trained in Study II. All the procedures described earlier were identical to those used in the current study.

The second training group, called the videotape-only training group, received the same training as described in the preceding experiment except that students were not provided worksheets on which to record behavioral indications of their cognitive responses to instructional stimuli that were presented during teaching. In order to teach the students in this group how to make the cognitive responses, overhead transparencies made from the same worksheets that were being used by students in worksheet-plus-videotape training group were shown. When the teacher on the videotape provided an instructional stimulus, the trainer simultaneously turned off the videotape and turned on the overhead projector. Thus, students in the videotape-only condition saw on the projection screen the worksheet that students in the Worksheet-plus-videotape group were using at their desks.

During the sessions in which students in the videotape-only group were being trained to make intended cognitive responses, the trainer used a pen to mark the transparency in the same manner as the behavioral indicators that were being taught to the students in the worksheet-plus-videotape group. In other words, the videotape-only training group received exactly the same instruction as did the worksheet-plus-videotape training group except that students in the videotape-only training group could not mark up worksheets. Instead, they were trained to carry out the mental equivalent of the behavioral indicators while looking at the projection of the worksheet. As the training sessions progressed, the trainer stopped turning off the videotape so that students in the videotape-only group were required to carry out their tasks in the same time available to students in the videotape-plus-worksheet group. In this group, then, students did everything that those in the worksheet-plus-videotape group did except write out behavioral indicators for intended cognitive responses. Otherwise, all aspects of the training in this group (e.g., answering test items, receiving feedback) were the same as

those for the videotape-plus-worksheet group.

The third group to receive training in this study did not view the videotapes of the lessons. They only were provided the same worksheets that were handed out in the worksheet-plus-videotape training group and were instructed in the intended cognitive response and behavioral indicators. Thus, this group received no instruction to identify instructional stimuli as cues about when to apply the intended cognitive response during lessons. In all other respects, this group was the same as the videotape-plus-worksheet group.

There were two groups used for comparison in this study. The first of these, labelled the videotape-only comparison group, merely watched the videotapes of the lessons, answered the test questions, and received feedback about their answers. They received no worksheets, nor were they given any instruction whatsoever about instructional stimuli, intended cognitive responses, or behavioral indicators. The second comparison group, called the videotape-plus-worksheet comparison group, were handed the same worksheets as described previously to accompany the videotaped lessons in which they were participating. However, no reference was made to instructional stimuli, intended cognitive responses, or behavioral indicators during the sessions. As well, none of the worksheets provided illustrative behavioral indicators of intended cognitive responses that were seen in the videotape-plus-worksheet condition. The videotape-plus-worksheet comparison group thus received exactly the same experience as the comparison group in Study II.

The range of times for training sessions and the mean time each group spent in sessions covering the first five sets of double segments about sleep or viewing the videotapes are presented in Table 7.

RESULTS AND DISCUSSION

In the analyses reported in the following sections, the original sample of 70 students is reduced to 44. The joint criteria used to select students who would be included in the

TABLE 7

Time Spent by Groups
Study III

Group	Range (min.)	Mean (min.)
Training		
Worksheet-plus-videotape	36-41	39
Videotape-only	37-40	39
Worksheet-only	39-46	41
Comparison		
Videotape-only	17-29	21
Videotape-plus-worksheet	34-39	36

analyses were the same as those for Study II with the added requirement that students who were receiving training to carry out intended cognitive responses had to be present for at least four of the training sessions. The analyses for this experiment parallel those for Study II.

Can Students Be Trained to Make Intended Cognitive Responses?

The worksheets used by students in the worksheet-only and worksheet-plus-videotape groups were scored for students' use of the consolidate and the compare-and-contrast responses. (Recall that the third group to receive training, videotape-only, did not use worksheets, hence they are excluded from all analyses of data regarding behavioral indicators of intended cognitive responses.) As was the case in Study II, students in the comparison groups who received worksheets were not expected to mark them in accordance with the behavioral indicators we defined, and this was the case.

The behavioral indicator scores for students in the two trained groups are displayed in Table 8. With respect to the written indicators of the consolidate intended cognitive response on the sleep curriculum, students in the worksheet-only group executed an average of more than seven out of eight components of the behavioral indicator. Students in the worksheet-plus-videotape group, while performing at a slightly lower level, nonetheless provided approximately 75% of the behavioral indicators of the consolidate intended cognitive response. In comparison to Study II, the lower mean for the worksheet-plus-videotape group is attributable to two students, one of whom scored only one (of 8) on this measure and another of whom scored only two. Overall, both groups performed well in providing the behavioral indicators for the compare-and-contrast intended cognitive response.

On the psychology curriculum, the worksheet-only training group performed to about the same level as did the worksheet-plus-videotape group in Study II on the consolidate behavioral indication score. The Study III worksheet-plus-videotape training group performed noticably less well than its counterpart in Study

TABLE 8

Means, Standard Deviations, and Maximum Scores for Behavioral Indications of Intended Cognitive Responses for Students Trained to Use Worksheets
Study III

Curriculum	Group	Consolidate	Compare-and-contrast
	Worksheet only	7.43	4.00
		.96	0.00
Sleep		(8)	(4)
	Worksheet-plus-videotape	5.90	3.60
		2.64	.84
	Worksheet only	33.71	23.29
		3.73	1.50
Psychology		(48)	(24)
	Worksheet-plus-videotape	27.40	23.20
		11.62	2.53

Note: Upper numbers are means, lower numbers are standard deviations. Numbers in parentheses are maximum scores.

II, although the lower mean can be attributed primarily to one student in this group who never provided behavioral indicators. If this student is eliminated, the mean for consolidate scores rises to 30.44. Both training groups performed equally well in providing the behavioral indicators for the compare-and-contrast intended cognitive response on the psychology curriculum.

Correlating the scores for behavioral indicators recorded during lessons on the sleep curriculum with those obtained during lessons on the psychology curriculum gauges the extent of transfer of the use of the cognitive strategies across the curricula. For the worksheet-only training group, the correlation for consolidate scores was -.05. The correlation for the compare-and-contrast scores cannot be calculated because of this group's zero variance for this score during the lessons on sleep. For the worksheet-plus-videotape training group, the correlation across curricula for the consolidate behavioral indicator scores was .80 ($p = .01$) and for the compare-and-contrast behavioral indicator scores the correlation was .67 ($p = .02$).

Unlike the earlier study, correlations between the consolidate and the compare-and-contrast behavioral indicator scores in the psychology curriculum were not negative. For the worksheet-only training group, this correlation was .58 ($p = .08$). For the worksheet-plus-videotape training group, this correlation was .22 ($p > .10$).

With the possible exception of a single student or two who failed to provide behavioral indications of intended congitive responses, these data show that students can be trained to carry out intended cognitive responses triggered by instructional stimuli. Moreover, the extent to which these students executed these behavioral indicators during the two curricula shows that students who were trained to make intended cognitive responses can carry them out during instruction. Transfer across the two curricula was very strong for the worksheet-plus-videotape training group, but near zero for the worksheet-only training group. We have neither data nor a theoretical argument to help us interpret this difference.

How Do Intended Cognitive Responses Relate to Achievement?

To investigate whether students' use of intended cognitive responses reflected by the written components of the instructional responses was related to their achievement, correlations were calculated between behavioral indicator cores and achievement scores (see Table 9). These correlations must be interpreted very cautiously owing to the small group sizes. Nonetheless, the data are quite clear in showing that consolidate intended cognitive responses evidence no statistically reliable ($p > .10$ for all correlations, two-tail tests) relationship to achievement across both curricula in either group. In contrast, the correlations between students' compare-and-contrast intended cognitive responses with the achievement measures on the sleep curriculum are generally strongly positive and usually statistically reliable despite small sample size in the worksheet-plus-videotape group. Correlations between compare-and-contrast intended cognitive responses and achievement could not be calculated for the worksheet-only group due to zero variance on the intended cognitive response measure. None of the other correlations relating to the psychology curriculum reached an acceptable level of statistical reliability ($p > .10$ for all correlations, two-tail tests). However, it should be remembered that these correlations are strongly affected by both small sample size and by outliers due to the few students who provided few or no behavioral indicators for consolidate intended cognitive responses during the psychology lessons.

The correlations were quite low between all measures of achievement on the psychology curriculum and students' use of intended cognitive responses reflected by behavioral indicators of the consolidate and the compare-and-contrast strategies. This same lack of relation was found in Study II. In Study II, however, there were moderately strong and statistically reliable correlations between behavioral indicators for <u>consolidate</u> instructional responses taken for the sleep curriculum and several indices of achievement from both curricula, whereas these relations were not replicated in Study III. Rather, the trained

TABLE 9

Correlations Among Behavioral Indicator Scores and Achievement for Students Trained to Use Worksheets
Study III

	Behavioral Indicators			
	Sleep		Psychology	
Achievement Measures	Consolidate	Compare-and-contrast	Consolidate	Compare-and-contrast
Sleep				
Multiple-choice	-31	-		
	-04	67*		
Attributes recalled	-09	-		
	-19	54*		
Compare-and-contrast	-06	-		
	-20	55*		
Total essay	-11	-		
	-25	50		
Psychology				
Multiple-choice	-26	-	22	09
	-24	18	04	-06
Attributes recalled	03	-	25	13
	-11	68*	-17	40
Compare-and-contrast	-16	-	34	02
	-30	57*	-19	08
Total essay	-06	-	32	09
	-14	71*	-06	27
Delayed Psychology				
Multiple choice	-46	-	-14	-13
	03	33	35	06
Attributes recalled	-10	-	34	11
	-13	53	-05	19
Compare-and-contrast	-27	-	24	-14
	-11	61*	03	25
Total essay	-11	-	30	09
	-05	59*	08	24

Note: Upper numbers are correlations for the Worksheet only group (N=7). Lower numbers are correlations for the Worksheet-plus-videotape group (N=10). Correlations could not be computed between sleep compare-and-contrast behavioral indicator scores and achievement for the Worksheet-only group due to zero variance on the behavioral indicator variable. Decimals are omitted on correlations.

* $p < .05$

group in Study III that was comparable to the trained group in Study II (videotape-plus-worksheet) evidenced moderately strong and statistically reliable correlations between behavioral indicators of the compare-and-contrast strategy taken during the sleep curriculum and achievement indices for both curricula. Unfortunately, the zero variance for the compare-and-contrast behavioral indicator scores taken during the sleep curriculum in Study II prohibits calculating this correlation in Study II and comparing the correlation to that observed in Study III.

It is clear from both Studies II and III that students are quite capable of acquiring both types of cognitive strategies. Furthermore, it is equally clear that using at least one strategy correlates with achievement. What is puzzling is why the correlation between use of the consolidate response strategy and achievement was unstable. Our speculation is that there are variables we have been unable to control or manipulate in these studies that influence this relationship.

The correlations indicating transfer of the response strategies from the sleep to the psychology curriculum were moderate to strong for both studies (.44 for consolidate in Study II, .80 for consolidate and .67 for compare-and-contrast in Study III). These correlations indicating transfer of response strategies would lead one to expect that correlations between use of the response strategies and achievement observed in the sleep curriculum also would be observed in the psychology curriculum. We have no explanation why this expectation was not achieved. While we believe that our methodology in Studies II and III has made progress in specifying components of the instructional environment that influence students' cognitive mediations of teaching and their subsequent achievement, we still lack a theory of sufficient explanatory power to account for the differences observed between the two studies.

Do Intended Cognitive Responses Enhance Learning from Instruction?

To examine relationships between training students to make intended cognitive responses and their achievement, several a priori comparisons were structured in the context of a backward selection regression procedure. In these analyses, as in Study

II, there were three major categories of predictors. The first category included only students' vocabulary scores, deviated about the grand mean. The second major category consisted of a set of a priori contrasts designed to provide information about differences between groups of particular interest. The contrasts are presented in Table 10. The first contrast examined differences between the two groups of students who viewed the videotapes without worksheets where one of the groups had been trained to make intended cognitive responses when presented with projections of worksheets. The second contrast specified differences between the two groups who watched the videotape and had worksheets, one of which had been trained to make intended cognitive responses and provide behavioral indicators on their worksheets. Contrast three defined the difference between two groups that were trained to make intended cognitive responses, but which differed in whether these cognitive responses also involved providing behavioral indicators. The difference between these two groups was operationalized as having worksheets upon which behavioral indicators could be made versus viewing a projection of the worksheet. The fourth contrast examined the effects of training students to make intended cognitive responses when they were provided only with the worksheets versus being able to watch the videotape and use the worksheet. Thus, contrasts one and two explored the relative contribution of training students to use intended cognitive responses within one training mode. Contrasts three and four compared the training components from Study II and tested the relative contribution each component made to the combined training condition.

The third major category of predictors considered in these analyses were aptitude-treatment interactions. These were created by multiplying each a priori contrast vector by the vector of vocabulary scores deviated about the grand mean.

As was the case in Study II, the vectors representing all these predictors were entered into the regression analysis at its first step. Thereafter, predictors that did not account for a statistically reliable portion of variance in the dependent variable ($p > .10$) were removed from the equation until only

TABLE 10

A Priori Contrasts Among Groups

	Untrained		Trained		
Contrast	Videotape only	Videotape + Worksheet	Worksheet Only	Videotape Only	Videotape + Worksheet
C1	-1	0	0	1	0
C2	0	-1	0	0	1
C3	0	0	0	-1	1
C4	0	0	-1	0	1

predictors meeting this criterion for statistical reliability ($p \leq .10$) remained.

Descriptive statistics for the vocabulary and achievement data from Study III are presented in Table 11. The results of the regression analyses and effect sizes for the a priori contrasts or aptitude-treatment interaction terms involving an a priori contrast (differences between adjusted means divided by the square root of the residual mean square from the regression analysis) are presented in Table 12. Group differences are reported if they describe effects that would be statistically reliable at least ten times per hundred comparisons ($p \leq .10$). As in Study II, this more than traditional allowance for type I error was made since this experiment was exploratory and involved very small group sizes.

The results of these analyses are quite varied. They are reported here organized according to the contrasts that were investigated and according to the order in which the achievement measures were collected, i.e., sleep, psychology, and delayed psychology achievement tests. Main effects for all of the contrasts are discussed first, followed by consideration of the aptitude-treatment interactions. As is discussed below, by far the most interesting and theoretically rich information is found in results involving aptitude-treatment interactions.

Table 13 summarizes the results of the 12 regression analyses reported in Table 12. It is organized by contrast and shows the dependent variables that were involved in reliable group differences, as indicated by the main effects portion of the table, and those that were dependent upon vocabulary by group aptitude-treatment interactions.

As one would expect, vocabulary scores were reliablly related to achievement test scores, contributing between five to 24% of the variance in the dependent variables. The main effect term for vocabulary score remained in 11 of the 12 equations, failing to be retained only in the equation where the psychology compare-and-contrast score served as the dependent variable. But even in this equation, vocabulary was associated with a statistically reliable aptitude-treatment interaction term.

TABLE 11

Means, Standard Deviations, and Correlations Among Vocabulary and Achievement Variables
Study III

	V control	V+W control	W only	V only	V+W	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Vocabulary	24.70[1] 4.85	24.67 3.64	22.43 5.86	19.63 4.07	24.70 5.56	79[2]	42	42	45	43	46	33	31	36	40	37	31	32
2. Sleep Multiple-choice	[illegible] [illegible]	5.89 1.17	5.00 2.24	4.75 2.49	6.60 1.65		72	39	39	44	23	14	08	19	37	12	16	16
3. Sleep Attributes Recalled	3.70 1.64	5.33 2.60	5.14 2.85	4.13 2.59	4.40 2.84			85	90	94	52	61	53	59	55	59	57	59
4. Sleep Compare-and-Contrast	6.60 1.65	7.22 2.68	6.71 3.35	5.88 2.53	5.80 3.61				77	97	45	60	63	67	45	60	70	69
5. Sleep Total Essay	11.20 3.58	13.67 5.72	12.00 6.06	10.25 4.68	10.10 6.44					82	48	60	56	64	50	61	66	67
6. Psychology Multiple-choice	14.60 3.24	17.33 2.45	14.86 2.79	15.13 2.90	17.50 2.88						54	45	44	43	80	45	44	46
7. Psychology Attributes Recalled	4.20 2.25	5.56 3.88	4.14 3.13	4.50 2.78	4.90 3.41							72	80	91	49	84	73	79
8. Psychology Compare-and Contrast	8.60 2.01	7.67 4.80	7.71 3.45	8.13 3.48	7.90 4.04								71	93	40	65	84	76
9. Psychology Total Essay	15.10 4.28	15.11 8.52	12.57 4.89	13.00 4.90	13.50 7.06									73	44	76	84	84
10. Delayed Multiple-choice	14.40 4.14	17.56 3.32	13.57 2.88	15.38 3.25	17.60 3.37										68	50	41	46
11. Delayed Attributes Recalled	4.10 2.47	5.11 3.48	3.57 2.82	4.25 3.41	4.80 3.29											76	79	90
12. Delayed Compare-and-Contrast	6.30 1.83	8.33 4.24	7.14 3.44	7.88 3.27	7.10 4.43												74	95
13. Delayed Total Essay	13.50 3.89	15.00 7.04	11.00 6.11	12.88 6.08	12.30 7.80													80

[1] Upper numbers are means; lower numbers are standard deviations. Means are unadjusted for vocabulary.

[2] Alpha reliability coefficients are displayed in the diagonal. Decimals are omitted in these and the correlations.

Note: V control = Videotape control; V+W control = videotape-plus-worksheet control; W only = Worksheet only; V only = Videotape only; V+W = Videotape plus worksheet

TABLE 12

Backward Selection Regression Analyses
Study III

Dependent Variable	Predictors	β	F/t[1]	p	% variance[2]	R
<u>Sleep</u>						
Multiple choice	Total equation		7.09	<.01	.30	
	Vocabulary	.23	1.67	.10	.05	
	C1	-.36	-2.70	.01	.12	-1.65
	Vocab x C2	.27	2.06	.05	.07	.27
	Constant	(5.99)	23.93	<.01		
Attributes Recalled	Total equation		5.76	<.01	.25	
	Vocabulary	.52	3.70	<.01	.24	
	C1	.30	2.07	.05	.07	-
	Vocab x C1	.28	2.08	.04	.08	.43
	Constant	(4.77)	13.92	<.01		
Compare-and-contrast	Total equation		7.90	<.01	.33	
	Vocabulary	.36	2.73	<.01	.12	
	C2	-.28	-2.17	.04	.07	-
	Vocab x C2	.40	2.92	<.01	.13	-.64
	Constant	(6.45)	18.92	<.01		
Essay Total	Total equation		6.34	<.01	.27	
	Vocabulary	.35	2.57	<.01	.11	
	C2	-.31	-2.31	.03	.09	-
	Vocab x C2	.32	2.28	.03	.11	-.79
	Constant	(11.45)	16.83	<.01		
<u>Psychology</u>						
Multiple choice	Total equation		6.21	<.01	.27	
	Vocabulary	.48	3.43	<.01	.20	
	C1	.35	2.35	.02	.09	.21
	C3	.27	1.80	.08	.06	.90
	Constant	(15.97)	40.61	<.01		
Attributes recalled	Total equation		5.05	.03	.09	
	Vocabulary	.33	2.25	.03	.11	
	Constant	(4.68)	10.71	<.01		
Compare-and-contrast	Total equation		6.91	.01	.12	
	Vocab x C2	.38	2.63	.01	.14	.33
	Constant	(8.01)	16.15	<.01		
Essay Total	Total equation		4.83	.01	.15	
	Vocabulary	.29	2.02	.05	.08	
	Vocab x C2	.25	1.75	.09	.06	.44
	Constant	(13.94)	16.69	<.01		

Dependent Variable	Predictors	β	F/t[1]	p	% variance[2]	E
Delayed Psychology						
Multiple choice	Total equation		5.38	<.01	.23	
	Vocabulary	.43	3.02	<.01	.16	
	C1	.23	1.67	.10	.05	.33
	C4	.27	2.03	.05	.07	1.08
	Constant	(15.77)	32.11	<.01		
Attributes recalled	Total equation		4.89	.01	.15	
	Vocabulary	.35	2.46	.02	.12	
	Vocab x C3	.24	1.71	.10	.06	.18
	Constant	(4.20)	9.62	<.01		
Compare-and-contrast	Total equation		5.04	.01	.16	
	Vocabulary	.27	1.93	.06	.07	
	Vocab x C4	.32	2.29	.03	.10	.17
	Constant	(7.63)	15.87	<.01		
Essay total	Total equation		5.47	.01	.17	
	Vocabulary	.28	2.03	.05	.08	
	Vocab x C4	.33	2.36	.02	.11	.17
	Constant	(12.75)	14.94	<.01		

[1] F-statistics are reported for the regression equation containing all predictors remaining in the equation ($p < .10$); t-statistics are reported for each predictor's slope coefficient.

[2] The figure reported for the Total equation is R^2 adjusted for shrinkage. Those associated with each predictor are squared part (or semipartial) correlations.

TABLE 13

Summary of Statistically Reliable Effects Involving Contrast Coded Vectors from Backward Selection Regression Analyses Study III

Contrast	Dependent Variable	Direction of Effect
Main Effects		
1	Sleep multiple choice	–
	Sleep attributes recalled	+
	Psychology multiple choice	+
	Delayed psychology multiple choice	+
2	Sleep compare-and-contrast	–
	Sleep total essay	–
3	Psychology multiple choice	+
4	Delayed Pychology multiple choice	+
ATI		
1	Sleep attributes recalled	+
2	Sleep multiple choice	+
	Sleep compare-and-contrast	+
	Sleep essay total	+
	Psychology compare-and-contrast	+
	Psychology essay total	+
3	Delayed Psychology attributes recalled	+
4	Delayed Psychology compare-and-contrast	+
	Delayed Psychology essay total	+

Note: The direction of effect for the aptitude-treatment interactions indicates the slope of the regression line for the trained group.

The first contrast compared the two groups that viewed the videotapes where one of the groups was trained to make intended cognitive responses and one was not. On three of the dependent variables (sleep attributes recalled, psychology multiple choice, and delayed psychology multiple choice), students trained to execute intended cognitive responses outperformed their untrained counterparts. The effect size statistics for these three dependent variables indicate that the average trained student performed at t e 67th, 68th and 63rd percentiles in contrast to the average student in the comparison group for these three dependent variables, respectively. However, on the sleep multiple-choice test, the comparison group considerably outperformed the trained group. The effect size was a large -1.65, indicating that the average trained student scored at the fifth percentile on this scale in contrast to the average student in the comparison group.

From these analyses, it would appear initially that training students to use intended cognitive responses on material projected by an overhead had deleterious effects on achievement, particularly with respect to cued recall on the multiple-choice test. There was a modest positive effect of this training on the number of attributes students recalled on the essay test, which is more akin to a free recall task. However, this relation vanished on subsequent tests measuring transfer to the psychology curriculum, where it was replaced by reliable, but modest, effects on both the psychology and delayed psychology multiple-choice tests.

The second contrast compared the effect of training students in intended cognitive responses by using both videotapes and worksheets with a group that used the same materials but was not trained to make intended cognitive responses. This contrast replicates the between-group comparison in Study II. Main effects for this contrast were found on two dependent variables, both of which were derived from the sleep essay test. On both sleep compare-and-contrast and sleep essay total scales, the comparison group had, on average, higher scores than did the trained group.

Both of these effects were fairly strong, with the average trained student scored at the 26th percentile in contrast to the average comparison group student on the sleep compare-and-contrast variable. The effect size for sleep essay total score placed the average trained student at the 21st percentile relative to the average comparison group student. The main effect for the contrast for these two groups did not appear in the analyses of the psychology or delayed psychology test data.

These data suggest that the training was deleterious again, at least with regard to its immediate effect. Some modest support from training emerged on the psychology and delayed psychology multiple-choice tests, but the support was not startling. However, the aptitude-treatment interaction analyses modify these interpretations considerably. Before turning to discuss these more complex findings, the data comparing the full training package (videotape-plus-worksheet) to the videotape-only and worksheet-only training groups are presented.

Contrast three compared the videotape-only group to the videotape-plus-worksheet group. Contrast four compared the worksheet-only group to the videotape-plus-worksheet group. There was an effect for the former on psychology multiple-choice and for the latter on delayed psychology multiple-choice scales. Both effects were somewhat strong, elevating the average student in the group receiving full training, in contrast to the average student in the respective comparison group, to the 81st percentile for contrast three and the 86th percentile for contrast four. Thus, there appears to be no support for the hypothesis that the worksheet and videotape components of the training are debilitating when used jointly.

There were nine statistically reliable aptitude-treatment interaction terms in the regression analyses. All indicate a positive relation between vocabulary score and achievement for the students trained to use intended cognitive responses, and a negligible or slightly negative relationship between vocabulary and achievement for the untrained students in the comparison groups.

Contrast one was involved with only one aptitude-treatment interaction term involving the number of attributes recalled in the sleep essay test. In this case the slope for the videotape-only treatment group was .48, and the slope for the videotape-only comparison group was .03.

Contrast two was involved in five aptitude-treatment interaction terms, by far the most of all the contrasts. These aptitude-treatment interactions indicate that students with higher verbal ability, as measured by the vocabulary test, had higher scores on the various achievement tests only if they were trained to make intended cognitive responses. The relationships were negative between verbal ability and achievement in the untrained group for all of the aptitude-treatment interactions involving contrast two. The range of raw regression weights for the trained group in these seven dependent variables was from .27 to .94, while the range for the comparison group was from -.08 to -.42.

Contrast three interacted with vocabulary on the scale measuring the number of attributes recalled on the delayed psychology test. The regression of the number of attributes recalled on vocabulary for the videotape-plus-worksheet trained group resulted in a regression coefficient of .42, while the comparable statistic for the Videotape-only trained group was -.01. Similar aptitude-treatment interaction terms were found for contrast four on the delayed psychology test. On the essay total score and on the compare-and-contrast score, the videotape-plus-worksheet students with higher vocabulary scores achieved more highly than both their lower ability peers in the same group and the students in the worksheet-only training group.

CONCLUSIONS TO THIS POINT

The results of Study I indicated that students were cognitively active as they responded to teaching in classrooms. The data showed that teachers' intentions for students' cognitive processing and students' reports of the cognitive responses they thought their teacher intended them to use were sometimes mismatched. As well, there were instances where these appeared to

be a high degree of correspondence between teachers' intentions and students' cognitive responses. We interpreted the findings from Study I to hold promise for the cognitive mediational model as a means for guiding research and suggested that the phenomena warranted experimental analysis. Studies II and III were designed, in part, to determine if elementary school students could be taught to respond to instrutional stimuli systematically with strategic cognitive operations.

Studies II and III provide abundant evidence that students can acquire cognitive strategies for learning in classrooms and that these strategies can be employed by students in a simulation of classroom tasks and conditions. Students in three groups in the two latter studies provided evidence that they used the intended cognitive responses that they had been taught in training sessions by tracing components of these strategies on the worksheets. All participants in the trained group in Study II and in the worksheet-only group in Study III obtained perfect scores on the behavioral indicator for the compare-and-contrast cognitive response for the sleep curriculum. The other group trained to use worksheets in Study III, videotape-plus-worksheet, had nearly perfect scores (M=3.60, maximum score of 4). For the consolidate response, the mean scores for these three groups ranged from 75% to 93% of the maxiumum score.

It was not difficult to train students to make intended cognitive responses during instruction that also could be manifested in observable performance. This finding is valuable for three reasons. First, that students did not rebel at the task of learning or using intended cognitive responses during teaching sustains the applicability of the cognitive mediational model as a potential guide for helping students to profit from teaching. Second, this finding bodes well for future research about the cognitive mediational model. Researchers will be able to design treatments that include students' cognitive mediations as independent variables in experiments by training these mediations first, and then verifying that students use them during teaching. This check on implementation validity is a necessary

step in creating valid explanations about whether and how teacher behaviors relate to students' achievement by influencing students' thinking during instruction (Doyle, 1978; Winne, 1982 a, b). Finally, the fact that students can manifest aspects of their cognitive activities during teaching can provide teachers with evidence about how well students are following the plan of a lesson from moment to moment. Such interctive data describing how students are cognitively manipulating content to be learned may become an important input to teachers' interactive decision making, supplementing information about what students have acquired at particular points in lessons (Marx & Winne, 1981).

In a sense, establishing whether students could be taught to use cognitive strategies was the easiest question to ask and to operationalize. More difficult questions can be raised that address the coordination of the two cognitive strategies by students as they learn from teaching, the transfer of the strategies over time from one curricular area to another, and the utility of the strategy for acquiring curriculum information during instruction and subsequently retrieving it in response to test questions. These latter questions were far less successfully answered by the latter two studies.

The best set of data for examining whether students' were able to use both strategies that they were taught to apply during the videotaped lessons was in the psychology curriculum. Because of the zero variance in two of the trained groups in the compare-and-contrast behavioral indicator for the sleep curriculum, correlations could not be computed that included the sleep compare-and-contrast behavioral indicator data. Usable data were available, however, for both behavioral indicators obtained during the psychology curriculum for all three groups in the two studies who were trained to make behavioral indications of cognitive responses.

In Study II, the correlation between the two behavioral indicator scores for psychology was unexpectedly -.52. Assuming that the behavioral indicator data reflect students' use of the cognitive responses in which they were trained, this correlation

shows that the students could not manage to use both responses during lessons as we had intended. But the correlations between use of both strategies in Study III were .22 for the videotape-plus-worksheet group and .58 for the worksheet-only group. Given that means were high for behavioral indicator scales during the sleep curriculum, some students under some conditions were quite capable of executing both cognitive strategies during both sets of lessons.

Several speculations can be offered about why the stability of use of the intended cognitive responses was not higher. First, although the students in the two studies attended very similar schools serving what appeared to be similar populations, there may have been individual differences among the students related to between-school factors across the studies that influenced the results. The vocabulary scores for groups in the studies suggest that the two samples were similar, at least in this respect. However, there may be individual differences arising from factors other than verbal ability that may have influenced students' acquisition and use of the instructional responses we trained. In paricular, some students may be more capable than others at perceiving teachers' intentions for cognitive processing.

We have demonstrated in another study (Marx, Winne, & Howard, 1982) that individual differences in locus of control and field independence were not related to students' ability to perceive cues that teachers use to signal their intentions for students' cognitive processes. However, both vocabulary and aptitude for perceiving instructional stimuli were related to students' achievement (see also Stayrook, Corno, & Winne, 1978). Both the Marx et al. (1982) study and Studies II and III reported here were based on videotape simulations of recitation teaching. Thus, it becomes doubly hazardous to generalize these unstable findings to classroom contexts. Research in naturally constituted classrooms is needed to explore these relations among aptitude for perceiving intended cognitive responses and achievement in more ecologically valid settings. The five related studies reported in the next chapter are one such extension.

A second possible source of instability in the findings of Studies II and III relates to the consistency in delivery of the training cross the two studies. Training for the videotape-plus-worksheet comparison group and the videotape-plus-worksheet trained group were delivered by different trainers in the two studies. While we attempted to control most of the parameters of training instruction to make them constant (e.g., by controlling time of exposure, directions, content students worked with, and comprehensively scripting the training activities), differences due to trainers may have occured. However, we have no data available to explore this possibility.

A third possible factor that might help explain the inconsistencies has to do with the behavioral indicators. It is necessary to verify the occurrance of the cognitive response if one wants to test hypotheses about cognitive events in classrooms. Experimental psychologists frequently use response times or measures about the organization of responses as data upon which to base speculations concerning inferred cognitive processes. Obtaining response times is extremely difficult in classroom research for obvious methodological and technical reasons. We made use of organization of responses through the scoring system for the essay tests. But the arguments put forth in Chapter I claim that more immediate, in vivo data are required to verify that the hypothetical cognitive processes, the intra-instructional cognitive events that are the putative cause of learning, actually occur. Relying only on retrieval data as represented by tests of cued or free recall only partially reflects aspects of the acquisition environment where interactions among delivery variables, acquisition processes, and retrieval processes take place. Yet, clearly, it is the acquisition environment and its interactions with other facets of instruction that are most central to a study of teaching effectiveness. Hence, we developed the methodological feature of the behavioral indicator to attempt to operationalize the occurence of students' cognitive processes as they occur during learning.

One of the unfortunate features of the behavioral indicator is that the learner must devote time and cognitive resources to

the task of creating it. It is possible that the investments needed to create the behavioral indicator interfere with or, at least, diminish resources devoted to processing inherent in carrying out the cognitive strategy per se. Put differently, our insistence on operationalizing students' on-the-spot cognitive response by requiring them to produce behavioral indicators may be counterproductive. Even though we have data from the training components of Studies II and III showing that students can acquire the cognitive response and use the behavioral indicator, hindsight raises doubts that they achieved high levels of automaticity of these processes as a result of the training they were given. Hence, when put into a situation where they were not guided specifically by a trainer, their performance deteriorated over time.

Finally, Studies II and III provide evidence that students' use of the cognitive strategies we have created can influence learning. The correlational evidence showed that more frequent use of the strategies during the sleep curriculum was associated with higher achievement on the psychology curriculum. This effect occured with the consolidate instructional response in Study II and the compare-and-contrast instructional response in Study III. Beyond noting the inconsistency across studies, we are unable to explain what caused the instability of relationships between use of the two instructional responses and subsequent achievement.

Several aptitude-treatment interaction terms emerged from the analyses of Study III data. The most frequent of these involved the contrast between the videotape-plus-worksheet conditions, one of which received training in cognitive responses, and one that did not. All of these interactions showed that students with higher vocabulary scores who were trained to use cognitive responses were more successful on the achievement tests than low verbal ability students, while the reverse was true in the comparison group. This finding is supportive of the cognitive mediational model, although one would hope that these aptitude-treatment interaction effects would be attenuated by increasing

the degree to which the low ability students could use the cognitive responses productively.

Of course, one could speculate that more robust effects of cognitive strategy usage on achievement did not emerge because the two cognitive strategies we trained were simply not very useful. While this interpretation must always be considered in research of this genre, it is not favored here because large bodies of literature on human learning show that the cognitive responses we trained students to apply to content have a solid history of promoting learning (e.g., see Doyle, 1980). Perhaps learning from simulations of interactive classroom teaching is radically different from the settings out of which these positive findings arise. Although we doubt this, too, the differences do need further specification (Snow, 1974). The next set of studies address some of these issues.

CHAPTER V

Training Students to Respond Strategically to Their Teacher's Instruction - Five Classroom Studies: Study IV

CHAPTER 5

Training Students to Respond Strategically to Their Teacher's Instruction: Five Classroom Studies

INTRODUCTION

The studies described to this point in our program of research provide three major conclusions about how the cognitive mediational paradigm relates to conceptions of teaching and to students' learning from teaching. The first study demonstrated that teachers and their students operate in ways that clearly reflect the mediating role of students' cognition in teachers' choices about instructional actions and in students' learning from teaching. Specifically, teachers attempt to influence the ways that students cognitively process information, and students respond, sometimes idiosyncratically, to perceived instructional events by engaging cognitive strategies that they activate either automatically or as a result of some amount of decision making.

The second major conclusion, arising out of the second and third studies, is that students can be trained to discriminate instructional stimuli from other events that occur during teaching, and to engage pre-arranged cognitive strategies upon a teacher's delivery of instructional stimuli. This provides evidence that teaching can do more than merely respond to the array of individual differences students display. Specifically, teaching can take some hand in shaping how students learn, as well as what they learn, by instructing students about links between instructional stimuli and cognitive strategies that can support learning. In effect, this validates the idea that teachers may facilitate students' learning by providing students with an enabling curriculum that consists of declarative and procedural knowledge about the cognitive activities involved in learning from teaching (Marx & Winne, 1981).

The third major conclusion, also emerging from the second and third studies, as well as from prior research, is that students'

achievement is partly a function of the cognitive strategies they activate in response to the instructional stimuli they perceive in teaching. Moreover, the results of Studies II and III demonstrate clearly that, under certain conditions, the cognitive strategies students bring to bear during teaching can influence learning deleteriously. Although we certainly did not plan to reduce students' achievement by training them to use cognitive strategies that were constructed from a foundation provided by hundreds of other studies in educational and instructional psychology, this did happen with some students. These negative findings, while perhaps explainable as due to a lack of automaticity in students' use of new cognitive responses to teaching, rather than pointing to a potential weakness in the foundational research, only reinforce further the need to continue research on teaching from the perspective of students' cognitive mediation of instruction. This is particularly important because of the repeated aptitude-treatment interactions showing that high verbal ability students were able to benefit from cognitive strategy training.

The five experiments that are reported in this chapter were undertaken to explore further the interrelations among students' use of cognitive strategies to learn from teaching and subsequent measures of their achievement. What distinguishes these experiments from the preceding pair are several features that help penetrate even further the black box of students' cognitions during everyday teaching.

By the time students enter the upper elementary grades, it is almost certain that they have developed a repertoire of cognitive responses to the instructional stimuli that appeared in the thousands of lessons in which they have participated earlier. It seems reasonable to assume that, like in all other areas of learning, students will display individual levels of achievement both in the range of cognitive responses they make to instructional stimuli, and in their developed abilities to identify instructional stimuli keyed to specific intended cognitive responses (Marx et al., 1982). In other words, we hypothesized that students have developed different levels of what

might be labeled "perceptual acuity" for intended cognitive responses. Students with different levels of this aptitude may respond differently to teaching, and this subsequently may influence their achievement. To explore this hypothesis, we devised a procedure to gauge students' aptitude for perceiving their teacher's intentions about students' cognitive responses during lessons. Scores generated by this procedure were treated as representations of an aptitude in our analyses of achievement data.

A second feature of this last quintet of experiments that sets them apart from Studies II and III is that we attempted to insure that students who were trained to make intended cognitive responses matched to an instructional stimulus did so more automatically than before. This goal was approached in several ways. Training sessions were longer and slightly greater in number to increase opportunity to learn the instructional response. To increase the relevance of training, the medium for training was videotapes of lessons in which the students had participated only minutes prior to the training session. To promote automaticity in transferring training from videotape representations of lessons to live teaching, we involved the students' respective teachers in the training. They retaught a brief section of the preceding lesson while students practiced identifying instructional stimuli, carrying out the associated intended cognitive responses, and writing down behavioral indicators of their cognitive responses.

A third feature of this final set of studies was the fact that they were carried out in an instructional context that was nearly identical to the context of these classrooms prior to our interventions. In fact, the only difference in the teaching that each of the five teachers delivered as a result of our work with them and their students was that, at times the teachers thought appropriate (and that had been discussed prior to the lesson), a simple instructional stimulus was delivered. In every other respect, the nature of their instructional delivery was left to their discretion, just as it would be had we not enlisted their

participation in our research. With respect to students' experiences, the only difference in the instructional context for untrained students was that we distributed booklets in which they recorded class notes on paper which had two vertical lines dividing the page into three vertical sections. All students were informed that we would collect these booklets. Since each of the teachers in our studies had involved their students in notetaking prior to our entry into their classes, the main differences for students was the fact that someone other than their teacher would be examining the booklets, and that there were lines on the pages of these booklets unlike "plain" notepaper. Untrained students were told to ignore these divisions because they were there only for the other group (trained students) in their class with whom we had been working. These minimal differences provided a naturalistic environment for the research that we judge to be nearly ecologically equivalent to that which students would normally experience in an average classroom in the participating school district.

It is possible, of course, that reactive effects of our work with the teachers occured, and that these effects changed the classroom environment in such a way that it had different effects on learning than had the environment prior to our influence. However, to a very large degree, the question of this influence is moot because it is impossible to conduct research on classrooms without influencing the environment in some way. We judge that it is possible to reduce the confounding nature of these influences by working with the participating teacher in a collaborative effort. In this way, research questions and methodologies used to investigate these questions can be tailored to the idiosyncracies of particular classrooms, while still having sufficient experimental rigor to speak to a theory of teaching that extends beyond the boundaries of individual classrooms. The research design used for the five experiments reported in this chapter was created to take advantage of collaboration with teachers in the service of enhancing the value of our intrusion for the teachers

who worked with us, and extending the theoretical generality of our research.

Overview of the Studies

All five studies reported in this section followed a common format consisting of a preparatory period, during which the research staff worked with the participating teachers to develop curriculum materials, and three subsequent phases constituting the experiments as such. Each study corresponds to a classroom that participated in the project. Because of platooning in this district (the same as for Studies II and III), one of the teachers, Mrs. Christy (a fictitious name used to insure confidentiality), taught three of the five classes of students. Mr. Dixon and Mr. Forrester (also fictitious names) each taught their own class. The nature of work undertaken in this study required that each teacher have constant and easy access to a member of the research team. Thus, at the first meeting that initiated the period during which curriculum materials were developed, each teacher was paired with a research assistant for the duration of the study.

During the preparatory period, the entire research staff worked with teachers to create curriculum materials and achievement tests that would be used in the three phases of the experiments. Three units of curriculum were developed, one for each experimental phase. The topics covered in the units were: the human sense of hearing, the sense of vision, and an introduction to local ecology. Although a large proportion of concepts and other aspects of these curriculum units were comaprable across the five classes, there were important idiosyncracies as well. This reflects our intent to conduct the research in as naturalistic a fashion as possible while insuring that our data would provide a foundation for valid interpretation.

Following these curriculum development activities, there were three phases to the study. The first phase, called the mapping phase, was designed to produce a comprehensive description of the intended cognitive responses characteristic of the natural teaching in each of the five classrooms. A four-lesson unit on

the sense of hearing served as the curriculum for the mapping phase. The second phase of the study, the training phase, was designed to train a randomly chosen half of the students in each classroom to use a generalized intended cognitive response when prompted by their teacher's instructional stimulus. This training phase was carried out when the lessons comprising the unit on the sense of vision were taught. In the third phase of the experiments, the implementation phase, all students in each classroom participated in recitation lessons about the ecology unit taught by their regular teacher. The purpose of this phase was to test whether the training that half the students had received in phase two would generalize to everyday teaching, and whether trained students' use of intended cognitive responses to their techer's instructional stimuli would affect their learning of information in the ecology unit. The nature of each phase is described in more detail in the following sections.

METHOD

Context: Mr. Forrester, Class A

Participants. The students in Mr. Forrester's class were 27 grade six children in a small elementary school. The school, which had one class at each of eight grade levels, is located in a middle socioeconomic suburban area consisting mainly of single family dwellings. Students in Mr. Forrester's class were predominantly white, though his class included approximately 20% Asian students. Of the 27 children in the class, 15 were girls and 12 were boys.

Class format. Students normally were seated in rows of four or five but this varied depending on the nature of the lesson to be taught. Students felt free to move their desks before a lesson, though this usually entailed moving closer or further from the board. Rarely did students exchange desks or alter their position in a row. Two or three students were not in rows but were seated in more private spaces in the class, either by a cabinet or near a corner of the room.

Mr. Forrester kept his desk at an angle to the class by the doorway to the hall. On one side of the class was a sliding screen separating Mr. Forrester's class from the adjoining grade five class. This screen prevented st dents from seeing activities in the other class, but noise from the grade five class occasionally interrupted Mr. Forrester's lessons. At the front of the classroom was a blackboard and lectern from which Mr. Forrester gave his lessons. Occasionally he would use the blackboard on the side wall, opposite the sliding screen. At the back of the class were several windows and a door to a playing field.

The classroom was filled with a wide array of resource materials including a collection of musical instruments and a micro-computer. Mr. Forrester had divided sections of the wall for posters and students' work relating to various aspects of the grade six curriculum. He also used charts to monitor student homework and other activities.

Mr. Forrester typically used an inductive approach to teaching. He would usually begin a lesson with a review of the last lesson and a brief overview of the upcoming lesson. He would then ask a series of questions to cover the points in his lesson. Mr. Forrester reacted encouragingly to students' responses while indicating the correctness of the response. With his guidance, more often than not the students eventually arrived at the intended learning outcome without Mr. Forrester having to state it outrightly. During this inductive process, Mr. Forrester permitted moderate background chatter and activity, never demanding absolute, undivided attention to himself. However, if the noise level became too high or the students' activity too obviously off-task, he meted out discipline quickly and efficiently.

Context: Mr. Dixon, Class B

Participants. The students in Mr. Dixon's class were 20 grade six children, ten boys and ten girls, in a large elementary school. The school is located in a middle socioeconomic suburban area consisting of single-family dwellings. The class was

predominantly white although there were about 30% Asian students. Mr. Dixon taught only science and social studies to this group. For the remaining subjects, this group was taught by another teacher.

Class format. During all but the last two sessions in the third phase of the study, the class sat in groups of four or five, each group at one of four large tables in the classroom. The desks of students in Mr. Dixon's homeroom class were spread around the perimeter of the room and were not used by the class participating in this study except during the last two sessions. For these lessons, the tables had been removed and the homeroom desks were put into conventional rows. The students sat where they wished. During the training phase of the experiment only, the group that received training sat at the two tables nearest the video equipment. During lessons, students remained at their tables except when sharpening pencils, using the washroom, or participating in a class activity.

Mr. Dixon used a table for a lectern and positioned it at the front of the classroom for easy accessibility to the blackboard. Frequently during lessons, Mr. Dixon used an overhead projector[1] aimed at a screen to the students' right. The classroom was essentially undecorated. Only a few maps and some examples of students' work were hung on the walls.

Mr. Dixon usually began a lesson with a review of the past day's lesson and a brief overview of the current lesson. He then would begin, using the blackboard or overhead projector, to introduce the new material. Mr. Dixon encouraged students' questions, often using them as starting points for a discussion of concepts specific to, or similar to but diverging from, those planned for a lesson. Because of the latitude he allowed in this latter respect, he did not always teach all the material he had planned.

Context: Mrs. Christy, Classes C, D, and E

Mrs. Christy taught three science classes, all of which participated in this study. They are referred to as Classes C,D, and E. Classes C and D were seventh grade classes of 12 girls and

ten boys and 12 girls and 11 boys, respectively. Class E was a sixth grade class containing 14 boys and 12 girls. The school was a large elementary school, situated in the suburbs and primarily populated by middle-class families. Students in Mrs. Christy's classes were predominantly white, though classes included a small proportion of Asian students, some of which participated in an English as a second language course which kept them from attending some of the lessons involved in this study.

Class format. Students in each class were normally seated in rows consisting of four or five pairs of desks each, but this varied depending on the activities for a lesson. On some occasions students changed desks, but this was usually the result of a disciplinary action by the teacher. Mrs. Christy was head teacher at the school and her lessons were frequently interrupted by the principal to discuss administrative matters.

Mrs. Christy planned her lessons in detail and referred often to her written plan during instruction. She primarily used an inductive approach to teaching. Lessons usually began with a review of the main concepts covered in the last lesson plus a brief overview of the upcoming lesson. Her lessons were mainly teacher-centered and she did most of the talking. As the lesson developed, however, Mrs. Christy began to engage students in discussion. She frequently used the board to draw diagrams and to write concepts that she wanted the students to copy into their notebooks. Although Mrs. Christy was friendly, she did not tolerate chatter, and on several occasions students were reprimanded and asked to leave the classroom if they could not refrain from talking. Seemingly because of this restriction, the students took opportunities to chatter when the teacher was not instructing.

Curriculum

The curriculum content for the three phases of the study was determined in a meeting involving all five researchers (the two principal investigators and three research assistants) and all three teachers. The criteria used to select content were:

1) the information should be new to the students so that measures of instructional effects would be sensitive to teaching that occured during the study rather than reflecting students' prior knowledge;
2) the content should fit the individual teachers' objectives and preferences as much as possible;
3) the material should be well-structured yet lend itself to the various kinds of testing, training and analysis that the study requires.

After careful consideration of these criteria, the teachers and researchers agreed that the subject areas that would be taught in the study would center around the science program for the upper intermediate grades, specifically dealing with the sense of hearing for the first phase of the study, the sense of vision for the second phase, and the science of ecology for the third phase.

The process for developing the science units following this meeting began by informally observing one or two science lessons in each class. During these observations, the amount of content (e.g., number of concepts) and duration of a typical science lesson were assessed. Based on this information and subsequent discussions with all three teachers, it was decided to plan four lessons in the unit about hearing, six lessons in the vision unit, and eight lessons in the ecology unit.

Following this preliminary work, the research team developed a general first draft of the curriculum that could serve as a stimulus to which the teachers could react. Based on this global plan for the curriculum, the teachers were asked to begin preparing an outline of more specific topics that would be covered in their lessons.

At a second meeting two weeks after the first meeting, the research team and all the teachers met for a full day to outline in more detail the major aspects of content for each of the three science units. During this meeting, each teacher and the research assistant assigned to him or her began to plan the overall structure of each unit and to list concepts to be dealt with in the lessons within those units. In addition to personal

knowledge, several sources were used including: Exploring Living Things: Laidlaw Exploring Science Program (1977); Stecher et al. (1976); Program on Teaching Effectiveness (1976); and several introductory psychology texts.

To provide for differences among the teachers in their objectives and preferences, the researchers and teachers had several subseuqent, often daily, meetings to add to and alter the agreed-on outline of concepts. The units and lessons that emerged from these adaptations are outlined in Appendix E. Since there was no intent to pool the results across the teachers' classes, these modifications to the general outline of the curriculum were expected and accepted in compliance with our second criterion for curriculum materials for the study.

Mapping

The goal for this first phase of the study was to produce a "map" of each teacher's characteristic intended cognitive responses. To accomplish this, a procedure based on that for the first year of the project was used to videotape each teacher's lessons and obtain his/her interpretations about intended cognitive responses in an interview following the videotaping. The procedure was as follows.

The teacher's research assistant set up the videotape equipment before class. Then s/he videotaped the lesson and took brief notes about incidents in the lesson that might contain intended cognitive responses. These notes and the corresponding segments of the videotaped lesson would be analyzed during the teacher interview. Criteria for identifying these events for later analysis were general. The research assistant had to be able to identify a possible cognitive response that students might use in response to an instructional event. The research assistants were prepared for this task through practice in a pilot study. After videotaping the lesson, the research assistant moved the equipment to another room and reviewed the notes to prepare for the interview that followed within an hour of the lesson.

The post-lesson interviews of the teacher ranged in length from 25 to 40 minutes and were audiotaped for further analysis. In the interview, the teacher viewed the videotape of the lesson.

and was asked to stop the videotape at points where s/he intended students to think in specific ways in response to particular teaching events. The teachers were prepared for this task prior to the first videotaping by a discussion of about ten minutes in which the research assistant described several intended cognitive responses identified in the earlier studies and provided several examples of each. However, if a teacher passed over a large portion of instructional events that had been noted by the research assistant during the lesson, the research assistant also stopped the videotape to obtain the teacher's analysis of these segments.

During the interview, teachers were told not to invent an analysis of intentions for students' cognitive responses when they had intended none during the lesson. The goal of the analysis was to obtain from the teacher as precise a description as possible about cognitive operations s/he intended students to use in response to his or her teaching. Sometimes, the teacher provided this description without prompting or requests for clarification from the research assistant. When this was not the case, however, the research assistant probed the teacher in a manner such that analyses were solicited from the teacher rather than the research assistant asking the teacher to verify one or another speculation. Only after non-directive probing proved insufficient did the research assistant request clarification in ways that directly structured the teacher's description of intentions for student cognitive processing. This procedure was used to reduce the chances that the researcher's biases would influence the teacher's self-reports. We have no data, however, to support the validity of this procedure, nor do we have any data which indicates that research assistants selected the same types of incidents for teachers to comment upon.

After a teacher was interviewed, the research assistant and both principal investigators analyzed the videotape of the lesson and corresponding audiotape from the interview. The goal of these analyses was to create categories of intended cognitive responses that were unique to each teacher. The procedure was as follows.

The videotape was stopped at points where the teacher had indicated in the interview that s/he wanted the students to engage in certain cognitive operations in response to his/her teaching. At these points, the intended cognitive response(s) described in the interview were written down, interpreted, and clarified based on findings from the first year's study. As these descriptions accumulated, preliminary category schemes were proposed and compared with previous and subsequent incidents within each lesson. These analyses were performed on each of the first three lessons on hearing in Mr. Forrester's and Mr. Dixon's classes. Because much of her teaching proved to be so similar across her three classes, we sampled only the first two of the four lessons on hearing in each of Mrs. Christy's classes (a total of six lessons) to reduce our demands on her time. These analyses produced an idiosyncratic "map" of the intended cognitive responses each teacher sought to have students use when s/he cued students by delivering instructional stimuli during teaching.

Following these analyses, tallies of each category of intended cognitive response were made along with annotations about their typical placement in lessons (e.g., in the introduction, after question-and-answer sessions). Eight intended cognitive responses used very frequently by each teacher were common among all three teachers. Thus, these eight categories of intended cognitive responses were used to describe teaching for all three teachers, although the frequency with which each teacher used specific intended congitive responses varied. These eight intended cognitive responses and the phrases we used to describe them to teachers are presented in Table 14.

After generating this list of indended cognitive responses, each research assistant verified with their teacher that this list accurately reflected the characteristics of their teaching in general, and that the specific instructional behaviors and cognitive operations they intended students to engage in response to these events were paired correctly. Each teacher agreed with our proposed list.

TABLE 14

Intended Cognitive Responses

(1) Monitor Comprehension:

Check to see if you understand.

(2) Compare Codes:

Think what is the same about the term and a diagram.

(3) Retrieve:

Try to remember something in particular.

(4) Monitor Other Students' Answers:

Listen carefully to what the next students say.

(5) Compare Attributes:

Figure out what is the same and different about the two ideas.

(6) Rehearse:

Say this idea once or twice in your mind.

(7) Orient to Microscopic Elements:

Pay close attention to the very next thing.

(8) Retrieve Attributes and Generalize:

Remember the key parts of an idea and use them to get an answer.

The fourth lesson in the unit on hearing was not used in the mapping phase. It was used to create aptitude tests that measured students' accuracy of perceiving their teacher's instructional stimuli and intended cognitive responses. These tests are described later in the section on instruments.

Generalized Intended Cognitive Response (ICR). The eight common intended cognitive responses identified by mapping the three teachers' instruction constituted too large a set to consider training students to execute all of them in response to eight distinct categories of instructional stimuli that a teacher might use. The problems we faced, then, were those of balance and economy. On the one hand, there were practical constraints during the training phase of the study regarding time and students' abilities to learn intended cognitive responses. On the other hand, we wanted to train students to use a large number of the intended cognitive responses their teacher cued in lessons so that their cognitive responses to teaching would be valid and would have an optimal chance to influence learning the ecology curriculum.

The solution adopted was to attempt to create a generalized intended cognitive response that had several properties. First, it should be a composite of as many as possible of the discrete intended cognitive responses identified in Table 14 to maximize its "content validity." Second, it should be theoretically applicable to several different forms of learning that would be tested by different types of items on the objective and essay achievement tests administered at the end of the ecology unit. This would increase its usefulness for students. Third, it should be amenable to being cued by a single instructional stimulus at quite different points in a lesson. This would reduce the demands placed on students and their teacher by requiring students to watch for only one instructional stimulus during lessons, and by requiring the teachers to work consciously at only one change in their teaching style. The need to minimize changes in the teacher's typical delivery also was necessary so that students not trained to use the generalized intended cognitive response would

not find their teacher's instruction substantially altered, thereby disrupting previously acquired cognitive activities for the comparison group during lessons. Fourth, the intended cognitive response needed to be structured so that students could provide instructionally unobtrusive but observable (written) indicators that they had executed the appropriate cognitive operations on the correct curricular information, i.e., produce a behavioral indicator of their cognitive responses to teaching.

The generalized intended cognitive response we produced was a multi-step cognitive plan for processing information. Each teacher signaled students to use this generalized intended cognitive response for important information by uttering a unique tag, the instructional stimulus, like "Make sure you ..." We label the pair formed by the teacher's instructional stimulus (IS) and the students' intended cognitive response (ICR) an <u>IS-ICR unit</u>. Several paths through this plan were possible, depending on the student's state of knowledge and whether the information had already been presented or would be presented shortly in the form of an answer to a teacher's question (see Figure 3). This cognitive plan incorporates all eight of the intended cognitive responses listed in Table 14, although a single execution of an IS-ICR unit rarely would involve all eight intended cognitive responses. This generalized IS-ICR unit met the first three criteria we required of an intended cognitive response.

To address the fourth criterion we demanded of an intended cognitive response, we sought to create a behavioral indicator of the generalized intended cognitive response. Four other criteria had to be met in creating this indicator. First, the behavioral indicator had to be simple so that students could learn it easily. Second, it had to require a small amount of time for its production so as not to interfere with the rapid flow of instruction in the classroom. Third, it should require that a student engage in the intended cognitive response in order to produce a correct behavioral indicator, i.e., be dependent on the intended cognitive response. Fourth, the behavioral indicator had to reflect the various parts of the generalized intended

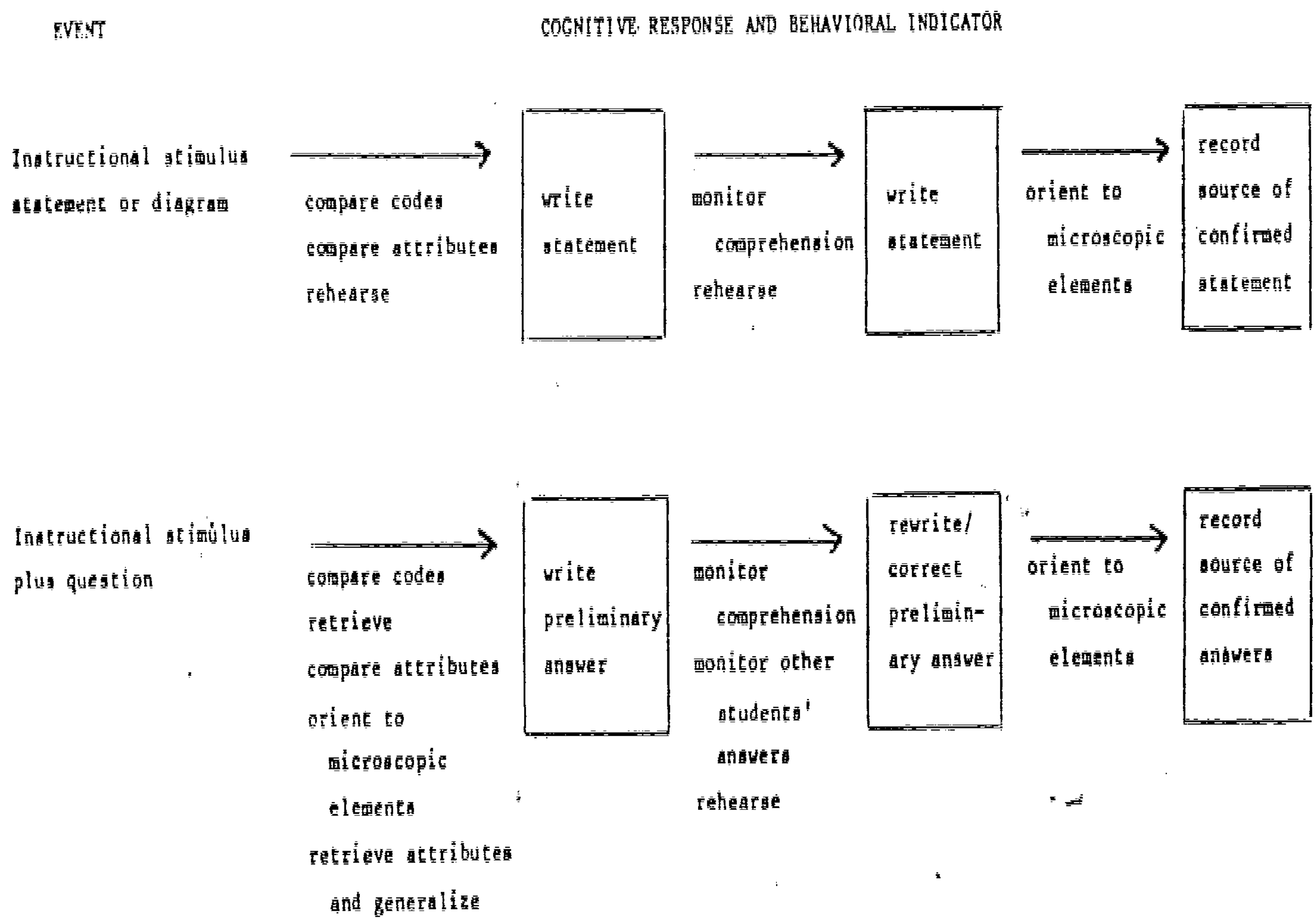

Figure 3. Correspondence of Cognitive Responses to Parts of Behavioral Indicator for Varying Instructional Events

cognitive response so that it represented the several covert, cognitive activities students engaged in response to instructional stimuli.

The behavioral indicator we designed to meet these requirements had three parts. Each part reflected the product of information processing corresponding to one step of the generalized intended cognitive response. The behavioral indicator for the first stage of the generalized intended cognitive response consisted of the student writing out a short but meaningful phrase. The phrase either repeated the teacher's statement, was a verbal representation of a diagram on the chalkboard, or represented that student's preliminary answer to the teacher's question. This observable act reflected intended cognitive responses noted in Table 14 in the following ways. Forming a preliminary answer to a teacher's question entailed the intended ive responses of orienting to microscopic elements of truction to identify appropriate content, retrieval of attributes plus generalization, or comparing attributes of concepts. The particular cognitive response depended on the kind of question the teacher had asked. Formulating a verbal expression about a diagram reflected the intended response of comparing codes. Writing a repetition of the teacher's question demands at least one rehearsal of that information.

To reflect the second stage of the generalized cognitive response, students also wrote a short phrase. It included information they had confirmed as important and correct. This component of the behavioral indicator for the generalized intended cognitive response could reflect several of the simple intended cognitive responses listed in Table 14. When students judged the correctness of the information that they had in mind before writing it, they were monitoring comprehension. An even more valid indication of students monitoring their comprehension was obtained when, in the context of a teacher's question, the phrase they wrote corresponding to the second stage of the generalized intended cognitive response differed from that written for the first stage. This suggests they were checking the accuracy of

information they processed during the lesson. As they wrote this information, it was rehearsed at least once. If the information being processed originated from another student, they had to monitor other students' answers before monitoring their comprehension and rehearsing.

In order to obtain more general evidence about students' patterns of attention to sources of information, we included a third part in the behavioral indicator. This part required students to record the source confirming that information they had cognitively processed in the second stage of the generalized intended cognitive response was important and correct. This confirmation could originate from any one of three sources: their own reasoning, information that the teacher stressed verbally or highlighted on the chalkboard (e.g., pointing at a part of a diagram, underlining a word), or another student's response that the teacher acknowledged as correct. Operationally, students wrote a code indicating the source of information. The codes were M for me (i.e., self was the source), T for teacher, B for the board, or S for another student. Thus, all eight of the intended cognitive responses in Table 14 that comprised the generalized intended cognitive response were reflected in one way or another in the three stages of the complete behavioral indicator.

Training

There were two purposes of training: to train students to use the generalized intended cognitive response when cued by their teacher's instructional stimulus, and to train them to provide behavioral indicators of the cognitive operations they carried out during lessons. For this latter purpose, students were provided with lined notebook paper that had been divided into three columns. The leftmost column was used to record the results of the first stage of the generalized intended cognitive responses. The rightmost column was the place for recording their responses representing the second stage of their cognitive processing, while the middle column was reserved for the codes (M, T, B, S) indicating the source to which they oriented to complete this second stage.

A training program consisting of six sessions was designed to introduce students and their teacher, who also attended these sessions, to the nature of the IS-ICR units and how these units would be operationalized in the upcoming lessons on ecology. The first four training sessions followed a common format. Each began with the research assistant instructing students about features of the generalized intended cognitive response and the corresponding behavioral indicators. To begin the session, the research assistant described what would be trained that day. Then, a pre-selected segment from the videotape of the preceding lesson on vision was shown. This was followed by the research assistant's description of how students were supposed to respond with the IS-ICR unit to that part of the lesson followed by a description of the corresponding parts of the behavioral indicator. Students then attempted these cognitive and behavioral activities, and were given feedback on their work by the research assistant. Three or four more opportunities for practice and feedback followed, each based on a new segment of the videotaped lesson. These activities occupied about the first two-thirds of each training session.

In the last third of the training session, the teacher retaught some of the material from the lesson that had been the basis for the preceding training activities. S/he practiced using the instructional stimulus and had an opportunity to observe the time students required to execute their mental and written tasks. Students practiced their instructional response during these periods to improve the skills they had worked on that day and in preceding training sessions. The research assistant provided corrective feedback to both the teacher and students about their performances.

The last two training sessions included only live practice with the teacher and students, and thus followed the format of the last third of the first four sessions. More detailed description of each session is provided next.

The first training period began with the research assistant providing a general description of the nature and purposes of training. Students were told that what they would learn in these

sessions should help them get higher marks on the achievement tests they would take on the vision unit and, subsequently, on the ecology unit. Next, the research assistant played the first segment of the videotaped lesson and identified the teacher's use of the instructional stimulus. All segments chosen for this first session illustrated a teacher's statement that implied an intended cognitive response. The research assistant described only the rehearsal component of the first stage of the generalized intended cognitive response by saying, "The teacher wants you to repeat that information several times in your mind." Then, the research assistant demonstrated on the board the first part of the behavioral indicator.

Practice examples followed in which the research assistant first asked students to verbalize what the IS-ICR unit was, and then to write the behavioral indicator of their cognitive processing in the leftmost column of their worksheets. Feedback about what the students wrote down was provided to insure that the information was correct, and to help students learn how to write brief but accurate descriptions rather than long verbatim accounts of the information they had rehearsed. Then, the live practice plus feedback session followed.

There were two objectives for the second training session. The first was to consolidate the training from the first session. The second objective was to introduce the second stage of the generalized intended cognitive response and the part of the behavioral indicator associated with it. The session began with the research assistant reviewing the previous day's session. Several segments of the videotape from the second lesson on vision that contained teacher statements tagged with the instructional stimulus were then used to practice these skills. As in the first session, the research assistant provided students with feedback about their work.

The research assistant then introduced students to procedures for using both the first and second stages of the generalized intended cognitive response when the teacher tagged a statement with the instructional stimulus. Specifically, they were told

that, following the first stage of the intended cognitive response, they were to rehearse this information at least one more time and then write it in the rightmost column of the worksheet. Practice segments from the videotape followed. The second training session ended with the teacher providing some live practice for students, supplemented with feedback from the research assistant.

With the two stages of the generalized intended cognitive response and their respective behavioral indicators now introduced and practiced for teacher statements, in the third training session, students moved on to extend the use of these skills to occasions where teachers tagged a question with the instructional stimulus. Using an example from the videotape of the lesson, students were told to: (1) try to answer the question, (2) write a brief answer, (3) seek confirmation of their answer, (4)mentally revise the first answer if needed, (5) check to be sure they understood the correct answer, (6) rehearse the correct answer several times, and (7) write out a brief version of the correct answer. For questions where students could not generate an answer for the question (step 1 above), they were instructed to draw a line in the leftmost column of their worksheet to indicate that they had noticed the instructional stimulus, and then proceed to the fifth step of this routine after seeking the correct answer to the question from another student or the teacher. A few practice plus feedback examples followed.

The only aspect of the generalized intended cognitive response remaining to be trained was for the student to indicate the source of the verified information that they had rehearsed and written on the worksheet. This also was an objective of the third training session. Students observed a segment of that day's videotaped lesson and received instruction from the research assistant regarding orientation either to their own cognitive responses, the teacher, the chalkboard, or other students' answers. They also were taught to record the code(s) on their worksheets that identified the source of the verified information they had rehearsed and written as a result of stage two of the

generalized intended cognitive response. A few further opportunities were provided to practice this new component in response to videotape examples. This training session ended with the teacher providing live practice for students plus feedback from the research assistant about the complete generalized intended cognitive responses and the corresponding behavioral indicators for both statements and questions.

The fifth and sixth training sessions provided the teacher and his/her students with further opportunities to engage in live practice and receive feedback from the research assistant. The objective of these sessions was to consolidate and sharpen the entire repertoire. The research assistant also continued to help students frame brief statements for the first and second parts of the behavioral indicator so that they would not fall behind the flow of information and events in regular lessons.

Aptitude Test

The model from which this research is derived postulates that, during instruction, teachers intend students to manipulate information cognitively in particular ways. Teachers cue these intended cognitive responses with instructional stimuli. The intended cognitive response (ICR) aptitude test was developed to gauge the extent to which students accurately perceived their teacher's intended cognitive responses for pre-selected instructional stimuli. A videotape of lesson four in the hearing unit served as the source of instructional stimuli and corresponding intended cognitive responses.

The ICR aptitude test was developed in several stages. Using the list of eight intended cognitive responses created by mapping the three teachers' instruction (see Table 14), we analyzed the fourth videotapes in each of the five classes to identify instances where an instructional stimulus cued intended cognitive responses. The goal sought in this analysis was to locate at least three salient illustrations in each lesson of each type of intended cognitive response from the list in Table 14. In total we identified 27, 23, 28, 31, and 32 such instances in each of the videotapes in the five classes A-E, respectively. Following our

preliminary analysis, the research assistant met with the teacher to validate the instances we had chosen. To do this, the teacher viewed the segments of videotape we had isolated with the list of the eight common intended cognitive responses from Table 14 in hand. The teacher's job was to select from the list the type of cognitive response that s/he had intended to use in that segment of the lesson.

The teacher's data from this validation, along with several other criteria we imposed that are described shortly, were the bases used to create a 20-item aptitude test for that class' lesson. All instances where we and the teacher had agreed on the intended cognitive response were sorted into the eight categories of intended cognitive responses listed in Table 14. The first criterion imposed on these instances in selecting 20 for inclusion on the ICR aptitude test was that the entire test had to contain two or three instances from each of the eight categories so that there would be a relatively balanced distribution of the types of simple intended cognitive responses over the 25-30 minutes of the videotape. Second, we chose instances in such a way that there were no long segments on the videotape, i.e., more than about 4 minutes, that did not contain an instance. Third, the instances were chosen so that the curricular information in the segments of videotape was as well-developed as possible. This decreased the chances that students' perception of intended cognitive responses would be influenced by weaknesses in their state of prior knowledge (see chapter 2). Fourth, all examples had to be of high technical quality, i.e., audio and video aspects had to be easily comprehensible. If there were more than three examples for each type of intended cognitive response, we reviewed our selections to choose the segment of the videotape which, in our judgment, most clearly represented each type and which met the foregoing criteria.

If there were fewer than two selections from a category of intended cognitive response, we reviewed the videotape to identify instances on which we and the teacher had disagreed in our original analyses. We then selected from among these instances

what was needed to complete the test in accordance with the foregoing criteria. In these few cases, ranging across the five classes from only 3 to 25 percent of all the instances that we had identified originally, we used the intended cognitive response that the teacher had selected as the "correct" intended cognitive response.

The selected examples of IS-ICR units formed a unique 20-item four-alternative multiple-choice test for each class. The stem of each item consisted of a segment of the videotape illustrating an IS-ICR unit in the teacher's list and a corresponding paper-and-pencil question, e.g., "How did Mrs. Christy want you to think now?" The three distractors listed for an item on the student's test paper were intended cognitve responses which corresponded to instructional stimuli other than the example portrayed on the videotape. They were chosen under the constraints that they had to seem to be plausible intended cognitive responses at that point, and that each intended cognitive response serve as a distractor with approximately equal frequency. The "correct" answer was the intended cognitive response that we and the teacher had agreed on independently. Each intended cognitive response occurred as a distractor approximately three times.

To administer the aptitude test, the research assistant showed the entire videotape of lesson four from the hearing unit to the students up to the point of the 20th IS-ICR unit. During playback, the videotape was stopped at those points which corresponded to a written item on the test. Each test item asked the students to indicte how they thought their teacher wanted them to be thinking at that point in the lesson. The videotape was restarted after all students had answered the item. The items contained in each of the aptitude tests are listed in Appendices F-H. The number of items students answered correctly served as their score on the ICR aptitude test.

Procedures for Scoring Notebooks

Students' notebooks in both the comparison and the trained groups were scored for the content recorded in them that related either to objective test items or to material for which points

were awarded in the essay items. For students in the trained groups, four types of scores were created. The first kind of score, called the behavioral indicator score, reflected the extent to which students' behavioral indicators of their cognitive operations were complete. This score is conceptually similar to the behavioral indicator scores in Studies II and III, varying only because the components of the intended cognitive response in this study differed from those of the earlier ones. Here a concept was identified for each test item on the objective test and for each point constituting the scheme for scoring essay items. The students' behavioral indicators were awarded from zero to three points for each concept depending on the number of aspects of the behavioral indicator that they completed on the notebook pages we had constructed for them. Thus, for the objective test on ecology, students in the trained group could get a maximum score of 72 (three points per complete behavioral indicator x 24 objective items). Separate behavioral indicator scores were generated for notes corresponding to items on the objective test and for notes corresponding to material that was awarded points in scoring the essay items. Students in comparison groups obviously could not receive a behavioral indicator score.

The second kind of score concerning students' cognitive operations during teaching measured the extent to which students took notes of any sort pertaining to material covered in the test items. This meant that any aspect of a student's notes referring to content covered in a test item was given a score of one. This score was also computed for students in the comparison groups. Thus, for the notes corresponding to items on the ecology objective test, students in both the trained and the comaprison groups could achieve a maxiumum raw score of 24. In order to allow comparison of this score with the behavioral indicator scores for students trained to make intended cognitive responses and provide behavioral indicators, this raw score was multiplied by three to place it on the same length scale as the behavioral indicator scores. This score, which awarded full credit for any component of a behavioral indicator for students in the trained groups, was called the total notes score.

The third and the fourth kinds of scores characterizing trained and comparison group students' notes taken during teaching reflected the extent to which they took notes about concepts that the teacher had tagged with an instructional stimulus, and about concepts which appeared on the test but which were not tagged by teacher's instructional stimulus. These are called the tagged notes score and the non-tagged notes score, respectively. In order to make these easily comparable to the foregoing scores, the sum of these two scores was fixed to place it on a scale with a length of zero to 72. Thus, the tagged notes score can be compared to the behavioral indicator score and to the total notes score by forming a ratio of the former to one or the other of the latter. The same type of comparison can be made for the non-tagged notes score. Also, the sum of the tagged and non-tagged notes scores equals the total notes scores. In the case of class A, the maximum score for the tagged notes scale was 54. The non-tagged notes scale had a maximum of 18 for this class. In the other four classes, the number of items on the objective test were equally divided between those tagged by teacher's instructional stimulus and those not tagged. Thus, the raw score maxima for these other classes were 36 and 36, respectively (12 items times three).

Scores for concepts recorded in students' notebooks corresponding to information that was awarded points when scoring the essay items varied from class to class. In class A, the maximum essay score was 72. In class B, it was 63. In classes C, D, and E, this maximum score was 38.

Achievement Tests

Because there was only general consensus among teachers as to the information that would be covered in each of the three curriculum units, teachers varied in the concepts and principles they taught. The research assistants met with their teachers to write items for the achievement tests for all three units prior to the beginning of instruction on each unit. These tests, comprised of objective items (multiple-choice, matching, fill-in) and short essay sections, were then reviewed and refined by the entire

research staff and then again by the teacher. Each test item was linked to one or more specific concepts in a lesson (see Appendices F, G, and H).

In order to assure content validity of the test items, the research assistant reviewed the videotapes and the notes taken during the lessons. If the information needed to answer an item was not covered, it was stricken from the test and a new item was generated covering content that had been taught. Prior to administering the achievement tests to students, all test items were checked a final time by the teacher for clarity and content validity.

The test for the hearing unit contained two essay items, and the tests for vision and ecology units each had three essays. Total essay score values across teachers ranged from 14 to 23 for hearing, 24 to 27 for vision and 21 to 23 for ecology. On the objective part of the achievement test, there were 12 items for hearing, 15 items for vision, and 24 items for ecology. Scores on the essay and multiple-choice tests covering the unit on vision are not used in analyses. These tests were included both for the teachers' purposes of assigning grades, and to acquaint students further with the nature of achievement measures they would experience at the end of the ecology unit. We also reasoned that these achievement scores could not tell much about students' learning vis a vis the cognitive mediational model because students in the trained group were in the process of altering their cognitive responses to teaching.

The achievement tests were created by sampling content from each of the lessons within a unit roughly equally. This criterion was only partially achieved because of a number of constraints. First, for the vision test, the number of items we sought was not divisible equally by the number of lessons that had been taught. This required us to include fewer items on the test from some lessons. Second, each lesson taught by the three teachers did not contain equivalent amounts of information. Some lessons contained more concepts and facts than did other lessons. Thus, a strict application of the equivalent item sampling criterion would have

required testing concepts and facts that were of minor importance because a particular lesson was conceptually and factually lean. Third, in some instances a teacher misrepresented the meaning of key information due to incomplete knowledge. We decided not to test students for information that they had been taught that was not factually correct. Thus, for some lessons, the pool of available content from which items could be sampled was reduced further. Fourth, on the ecology test, we wanted to distribute items on the test over both tagged information (i.e., information the teacher had tagged with the instructional stimulus) and untagged information (i.e., information not tagged with the instructional stimulus). The complete specifications for all of the achievement tests for the five classes can be found in Appendices F, G, and H.

All questions on both the essay and objective parts of the achievement tests were randomly ordered. For all tests, the essay items were completed before the objective part. There were four possible responses in multiple-choice questions which were randomly ordered with the proviso that there were an approximtely equal number of correct responses in any given position.

Scoring manuals were created (see Appendices F-H) to score the essays. Points were assigned according to the degree of correspondence between a student's response and the criteria established as acceptable in the manual. Each research assistant marked all the tests for the students in his/her assigned classes.

RESULTS AND DISCUSSION

For all of the analyses reported in this section, the original sizes of the training and comparison groups were reduced to the levels as shown in Table 15. Criteria used to make these reductions in sample sizes were as follows: Students could not be absent for more than one of the lessons in the hearing unit or the ecology unit, and had to be present for the ICR aptitude test, the objective and the essay tests on the hearing lessons, and the objective and the essay tests about ecology. In addition,

TABLE 15

Sizes of Groups Providing Scores Used in Analyses
Study IV

Class	Trained	Comparison
A	10	11
B	8	7
C	10	6
D	11	8
E	11	14

students in the trained groups had to be present for at least four of the six training sessions. Thus, in classes A, B, and E, students were allowed to miss two of the training sessions and still be included in the experiment provided that they met the preceding criteria. In classes C and D, where one training session of the six planned sessions had to be cancelled due to problems in scheduling, students had to be present for four of the five training sessions. These criteria ensured that students had an opportunity to learn the material presented in the lessons, and that there were scores available for them on the ICR aptitude test and prior achievement tests on the hearing unit that could serve as predictors in backward selection regressions analyses.

For at least one variable in each of the five classes, it was computationally impossible to produce internal consistency estimates of reliability due to problems of indetermina Moreover, some of the remaining estimates of reliability were less than classical test theory allows when judged against the correlations between variables, i.e., squared correlations exceeded the reliability coefficient for one or both variables. We conclude from these conditions that many variables measured in these classrooms were not internally consistent (heterogeneous; Cronbach alpha ranged from .05 to .79; Md = .60) and that the small sample sizes per classroom exacerbated this situation. Therefore, we have not reported reliability coefficients for variables in this section. However, on the basis of both correlations among variables and the regression analyses presented later in this section, we do not judge that the data is unreliable. Rather, we take the view that real teaching situations in the classrooms where we worked produced variables for which internal consistency estimates of reliability are inappropriate.

Can Students Be Trained to Make Intended Cognitive Responses?

Table 16 provides means, standard deviations, and t-tests among means for the trained and comparison groups in all five classes on the scores derived from their notebooks. Only for the essay-related scores for Class B were there statistically reliable

TABLE 16

Means, Standard Deviations, and t-Tests for Scores from Notebooks
Study IV

		Total Notes			Tagged Notes			Non-tagged Notes			Behavioral Indicator		
Class	Test	M	sd	t	M	sd	t	M	sd	t	M	sd	t[3]
A	Objective[1]	37.50[2]	14.23	.62	24.90	9.80	.43	12.60	4.86	1.00	30.80	15.58	2.60*
		33.55	15.05		22.91	11.48		10.64	4.11		-	-	
	Essay (72)[4]	20.40	4.43	-.93	14.10	3.18	-1.34	6.30	2.21	.23	15.00	6.45	2.58*
		23.73	10.52		17.73	7.99		6.00	3.55		-	-	
B	Objective	24.38	20.87	-.63	12.00	10.52	-.74	12.38	10.81	-.47	21.75	20.62	2.54*
		30.00	11.87		15.43	6.58		14.57	6.35		-	-	
	Essay (63)	7.13	6.98	-4.05*	1.00	1.41	-4.67*	6.13	5.87	-3.69*	5.63	5.91	1.98†
		22.29	7.52		4.00	1.00		18.29	6.90		-	-	
C	Objective	41.10	15.50	.51	17.40	8.34	.31	23.70	9.21	.62	31.30	11.74	5.37*
		36.50	20.21		16.00	9.61		20.50	11.13		-	-	
	Essay (48)	31.20	9.30	.77	12.60	2.15	2.30*	18.60	3.41	.16	16.80	6.83	6.63*
		26.50	15.23		8.50	2.71		18.00	8.90		-	-	
D	Objective	48.55	11.29	.98	26.74	4.13	.51	21.82	7.95	1.09	25.18	13.75	9.05*
		43.88	8.48		25.88	2.75		18.00	6.99		-	-	
	Essay (48)	45.00	4.45	1.46	36.00	4.45	1.46	9.00	0.00	0.00	19.73	5.66	23.98*
		41.63	5.66		32.63	5.66		9.00	0.00		-	-	
E	Objective	43.91	8.51	.81	19.36	4.91	.54	24.55	5.66	.85	25.64	10.42	7.25*
		40.93	9.68		18.21	5.58		22.71	5.10		-	-	
	Essay (48)	37.36	6.70	.35	13.91	5.89	-.96	23.45	2.95	1.51	17.82	7.44	12.06*
		36.43	7.05		15.88	4.31		20.57	5.75		-	-	

1 The maximum total notes score on the objective items was the same in all classes, 72.

2 Upper figures are for the trained group; lower figures are for the comparison group.

3 This t-test compares the behavioral indicator score to the total notes score within the trained group.

4 Numbers in parentheses are maximum total notes scores.

* $p \leq .05$

† $p \leq .10$

differences between the trained and comparison groups in their notetaking during lessons. In this single instance, students in the comparison group took more notes relating to concepts appearing as test items than did students in the trained groups. Thus, to the extent that students' notes indicate that some kind of cognitive processing of test-related information is taking place, there were no differences in the overall levels of cognitive processing exhibited by students in the trained groups and those in the comparison groups, with the exception of class B.

Also provided in Table 16 is a within-group comparison of the scores earned by students in the trained groups on the behavioral indicator score and the total notes score for concepts appearing on the objective and on the essay tests. In every case, students' total notes scores exceeded ($p < .10$) their behavioral indicator scores at statistically reliable levels. Thus, while students in the trained groups generally did not differ from their untrained counterparts on the number of occasions in which they engaged cognitions regarding curriculum content to be tested, their notes revealed that not all of these occasions completely reflected the three-stage generalized intended cognitive response they had been taught. Whether this difference indicates that students in the trained groups had internalized more of the intended cognitive response, thereby making the behavioral indicators perhaps seem less necessary to them, or whether some aspects of the generalized intended cognitive response were not carried out cannot be determined from these data. All that can be said is that there are some statistically reliable differences in trained students' execution of the generalized intended cognitive response for concepts appearing on the tests when compared to their total notes scores.

One way to measure the extent to which these differences occur is to calculate effect sizes. These effect sizes, calculated by subtracting the mean of the behavioral indicator score from the mean of the total notes score and dividing by the standard deviation of the total notes scores were as follows: in class A, .47 for the objective test concepts and 1.21 for essay

concepts; in class B, .13 for the objective concepts and .21 for the essay concepts; in class C, .63 for the objective concepts and 1.54 for the essay concepts; in class D, 2.06 for the objective concepts and 5.67 for the essay concepts; and in class E, 2.15 for the objective concepts and 2.15 for the essay concepts. Thus, with the exception of class B, there were fairly substantial differences in the extent to which trained students provided full behavioral indicators for the generalized intended cognitive responses they carried out corresponding to information that appeared on the tests.

Another feature of the data reported in Table 16 that warrants mention is the relatively large variance in several of the distributions of scores. Examination of frequency counts for scores revealed that the distributions of these scores were occasionally extremely skewed, with some students taking absolutely no notes whatsoever and other students taking relatively massive amounts of notes. Thus, even in the trained groups, there was wide variation in the extent to which the teacher influenced students to cognitively process information that appeared on the test. The fact that the trained students did not generally differ from comparison students in the number of notes they took on concepts that the teacher had tagged with an instructional stimulus might be interpreted to indicate that students in the trained groups missed many opportunities to engage in the generalized intended cognitive response that they had been trained to use. It also could be interpreted, however, to reflect the demanding pace of regular classroom instruction which might prevent students from providing complete indications of all the cognitive processing they undertook. In particular, it might be that students do not take notes on all concepts that are tagged, but take notes only on those concepts that they believe are particularly difficult or that warrant an "extra" amount of cognitive processing. As well, these data might indicate that we have still not overcome the methodological problems created by the use of behavioral indicators of cognitive processing. That is, the behavioral indicator may be interfering with the students'

cognitive responses. In the face of conflicting demands regarding whether they should complete the behavioral indicator or continue to execute the cognitive strategy, students may have chosen the latter over the former.

Overall, then, it can be said that students in the trained groups and the comparison groups took roughly the same quantity of notes pertaining to information that appeared on the tests. Thus, any differences that are observed on test scores cannot be attributed to the quantity of notes that were taken. Instead, differences in achievement scores must reflect qualitative features of the cognitive operations that the students undertook and reflected in the notes that recorded aspects of these cognitive operations during teaching.

How Do Intended Cognitive Responses Relate to Achievement?

To explore relationships between students' cognitive processing of content that would appear on tests and their test scores, correlations between the various scores derived from students' notebooks and achievement tests were computed. These are reported in Table 17 for each of the five classes. As can be seen in Table 17, there is wide variation in the degree to which students' quantity of notes is associated with their achievement. While statistical tests of these correlations are hampered due to the small sample sizes, with only several exceptions, the general finding is that students' notetaking, and by inference, their cognitive processing of information during teaching, is positively associated with their achievement. This seems particularly true for students who were trained to make generalized intended cognitive responses and provide behavioral indication when the objective tests served as the measure of achievement. In these cases, the median correlation was .48, whereas it was only .23 when essay tests served as the measure of achievement. This pattern was reversed for the students who participated in the comparison groups. For those groups, the median correlation between quantity of notes and achievement on the objective test items was only .23, while the comparable median correlation for the essay test items was .42.

TABLE 17

Correlations Between Achievement Scores and Notes Scores
For All Classes
Study IV

Scales Correlated	A	B	C	D	E
Objective					
Total scale and	44	68*	50	44*	63*
Total notes	-08	60	59	-15	42
Tagged items and	43	77	44	53*	40
Tagged notes	-03	45	52	-50	42
Non-tagged items and	21	61*	56*	49	26
Non-tagged notes	-14	23	75*	-13	16
Total scale and	34	70*	46	42	70*
Behavioral Indicator	-	-	-	-	-
Essay					
Total scale and	-23	36	19	61*	31
Total notes	43	74*	59	09	40†
Tagged scale and	-42	33	19	54*	43†
Tagged notes	11	71*	17	20	-02
Non-tagged scale and	13	13	-02	-[a]	-08
Non-tagged notes	84*	53	51	-	47
Total scale and	-02	33	-01	65*	64*
Behavioral Indicator	-	-	-	-	-

Note: Upper correlations are within the trained group; lower correlations are within the comparison group. Decimals are omitted.

[a] There was no variance on the Non-tagged notes for essay content in Class D.

* p <.05

† p <.10

Do Intended Cognitive Responses Enhance Learning from Instruction?

To test the relative influence that training students to make intended cognitive responses had on their achievement, comparisons were made to their untrained counterparts. The method of analysis here paralleled that for Studies II and III. In particular, backward selection multiple-regression procedures employing four major categories of predictor variables were applied to the data. The first category consisted of aptitudes: students' ICR aptitude test scores, their scores on a prior measure of achievement based on objective items from tests on the hearing lessons, and their scores on essay measures of achievement based on the hearing lessons. The second major category of variables included a measure of students' notetaking during lessons that corresponded to the dependent variable, i.e., tagged objective items, non-tagged objective items, tagged essay items, and non-tagged essay items). For instance, when tagged objective items was the dependent variable, the tagged notes score was used as the measure of classroom process from among the various notetaking scores that were described previously. The third major category was a single variable consisting of an effects coded vector comparing the trained (+1) versus the comparison group (-1) of students. The fourth major category of variables included seven interaction terms. One was a notetaking by treatment interaction, using the notetaking score that was matched to the particular dependent variable, as in the second category of predictors described previously. Three were formed by multiplying the three aptitude test scores from the first category of predictors by the treatment contrast vector. The remaining three were three-way interactions representing the product of aptitude x notetaking x treatment.

Descriptive statistics and the results of backward regression analyses are reported in Tables 18 through 32. The results are described first in terms of each of the five classes individually. An integration of the results across the five classes is then presented and reflected in Table 33.

A Statistical Note

Before presenting the results, three aspects of these analyses warrant some elaboration so that the information and statistics to be presented later can be understood more easily. The first item to be addressed is the method for interpreting a multiple regression equation that contains two or more terms involving either a contrast between the experimental and comparison groups, aptitude-treatment interactions, notetaking-treatment interactions, or all these types of predictors. A second issue that is discussed pertains to interpreting aptitude-notetaking x treatment interactions. The third item relates to the fact that only one measure of effect size can be reported per analysis even though two or more terms involving the contrast between the trained and comparison groups may appear in the results of an analysis.

Often a multiple regression analysis reported in the following sections contains several predictors involving the contrast between the trained and the comparison groups. To take a generalizable case, suppose that the analysis of the total objective scale (Y) had retained these terms: ICR Aptitude (X_1), Group (X_2), ICR Aptitude x Group (X_3), and Prior Objective Achievement x Group (X_4), plus a constant term (C) corresponding to the intercept. The multiple-predictor regression equation for this fictitious illustration could be written symbolically like this:

(1) $Y' = C + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4$

To explain how individual terms in this equation can be interpreted (keeping in mind the caveat in footnote 1, p. 106), it is advantageous to write equation (1) in a novel format, replacing the mathematical X_i with the words describing the predictors to which they refer and labeling each term in the equation. The rewritten version of the fictitious example would look like this:

(2)	Predicted Total Objective Scale = Constant	(a)
	+ b_1 (ICR Aptitude)	(b)
	+ b_2 (Group)	(c)
	+ b_3 (ICR Aptitude x Group)	(d)
	+ b_4 (Prior Objective Achievement x Group)	(e)

The principle that is applied to interpret any one predictor is to rewrite the equation to create two groups of predictors. One of these groups will be made up of all the terms in the equation that include the predictor that is going to be interpreted. For example, to interpret the effect of ICR Aptitude, terms b and d are grouped together:

(3) b_1 (ICR Aptitude) + b_3 (ICR Aptitude x Group).

This grouping of terms can be rewritten like this:

(4) $[b_1 + b_3(\text{Group})]$ ICR Aptitude

because the predictor ICR Aptitude was included in both of the original terms labeled b and d in equation (2). Now, when the code assigned to a particular group, either +1 for the trained group or -1 for the comparison group, is substituted into expression 4 for the word "Group," and when the numerical values of the partial raw regression coefficients are substituted for b_1 and b_3, a single numerical value for the terms in the brackets of expression (4) can be calculated. This single value is the partial raw regression coefficient for the predictor ICR Aptitude that is associated with the group whose code was substituted in the calculations. Hence, there are two regression coefficients produced by doing these substitutions twice, one for the trained group when +1 is substituted for Group, and another for the comparison group when -1 is substituted for Group. Now, we turn to the second collection of terms mentioned in the principle stated earlier.

All of the remaining terms in equation (2), labeled a, c, and e in that equation, are added together to form the second group of predictors referred to earlier:

(5) Constant + b_2(Group) + b_4 (Prior Objective Achievement x group).

If we combine expressions (4) and (5), they look like this:

(6) $Y' = [\text{Constant} + b_2(\text{Group}) + b_4(\text{Prior Objective Achievement x Group})] + [b_2 + b_3(\text{Group})]$ ICR Aptitude.

When it is noticed that values can be substituted for every one of the terms found within brackets, simply adding up these values to get one number makes this expression look suspiciously like a

simple regression equation. It involves only one predictor, ICR Aptitude; one regression coefficient, the term $[b_1 + b_3(\text{Group})]$; and a constant that is made up of several items added together, namely the terms in expression 5. The question that remains to be answered is how to choose values for these terms to produce one numerical value for the term that is the regression coefficient, and one for the term that is the constant.

Consider the term for the regression coefficient first. Numerical values for b_2 and b_3 are part of the results provided directly by the original multiple regression analysis. There are two numerical values for the word "Group"; +1 and -1. Thus, looking at $[b_1 + b_3 (\text{Group})]$ when +1 is substituted for Group, the numerical values for b_1 and b_3 are also substituted, and the arithmetic of multiplying b_3 by +1 and adding the result to b_1 is done, the end result is a regression coefficient for the training group. Repeating the procedure but substituting -1 for the word "Group" produces a second regression coefficient for the comparison group. Thus, there are really two different regression coefficients, one for each of the two groups symbolized by the word "Group" in equation (6).

Turning to the collection of terms inside the brackets representing the constant (i.e., expression 5), the word "Group" can be replaced by its code, i.e., ±1 for the trained group or -1 for the comparison group depending on which group we are working with for the moment. Since we are now describing a complete group of students, it seems reasonable to replace the words "Prior Objective Achievement" with a statistic that describes the group being considered. The number chosen for this substitution is a group's mean on Prior Objective Achievement. Next, substituting the values produced by the multiple regression analysis for Constant, b_2, and b_4, and doing the arithmetic yields a single numerical value. This value is the y-intercept for the group whose code was substituted for the word "Group" and whose mean score was substituted for Prior Objective Achievement. Thus, there will be two y-intercepts: one for the trained group and another for the comparison group.

The overall result of these operations is to produce two separate regression equations. One predicts Total Objective scores for students in the trained group. The other predicts these scores for students in the comparison group. Each regression equation has a simple form, namely:

(7) $Y' = c + bX$

Because the regression equations for each group have different regression coefficients, these two equations describe an aptitude-treatment interaction. Following the same procedures as just outlined, the aptitude—treatment interaction involving Prior Objective Achievement also could be described as two simple regression equations. In general, any predictor can be described similarly by following this procedure.

The next aspect of the multiple regression analyses that needs explanation is how to interpret aptitude-aptitude-treatment interactions or aptitude-notetaking-treatment interactions. In general, the same procedure as was just described is applied to these more complicated terms except that one more step is needed. Using the same example as before (see equations 1 and 2), we add on one more term involving an aptitude-notetaking treatment interction (X_5) where the notetaking variable is the Total Notes score on objective items. This produces parallels to equations (1) and (2) that look like this:

(8) $Y' = C + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5$

and

(9) Predicted Total Objective Score = Constant (a)
+ b_1 (ICR Aptitude) (b)
+ b_2 (Group) (c)
+ b_3 (ICR Aptitude x Group) (d)
+ b_4 (Prior Objective Achievement x Group) (e)
+ b_5 (Prior Objective Achievement x Total Notes x Group) (f)

Applying the principle of regrouping to the predictor Prior Objective Achievement, the following two groups of terms are created:

(10) Constant + b_1(ICR Aptitude) + b_2(Group) + b_3(ICR Aptitude x Group

and

(11) b_4(Prior Objective Achievement x Group) x b_5(Prior Objective Achievement x Total Notes x Group).

The terms in expression (11) can be rearranged to isolate Prior Objective Achievement like this:

(12) [b_4(Group) [illegible] Total Notes x Group)] [Prior Objective Achievement].

As before, substituting the codes for either the training group or comparison group in expression 10 and substituting the corresponding mean score for each group's ECR Aptitude Test produces a y-intercept for each group. The same procedure applied to the variables named in expression 12 produces a unique regression coefficient for each of the groups. Putting these two results together yields one simple regression equation for each group that describes the relationship between a student's predicted Total Objective Score and his or her Prior Objective Achievement score.

There is one elaboration of this procedure that now needs to be pointed out. Each time a numerical value was substituted for a predictor that was a continuous variable, such as the Total Notes score, the value chosen for substitution was the group's mean score. However, this particular choice is not the only value that could be used. For instance, a score equal to the group mean plus one standard deviation might be used instead. If this alternative value for the Total Notes score were substituted in expression 11, a different number would be created for the group's regression coefficient. In other words, within a single group of students, say the trained group, there is a different slope relating the predicted dependent variable, Total Objective Scores, and the aptitude, Prior Objective Achievement, for every score that can be recorded for the notetaking variable, Total Notes. Thus, the relationship between predicted Total Objective Scores and students' Prior Objective Achievement changes as scores on Total Notes change.

Conceptually, this is what interaction means. When there is a "simple" aptitude-treatment interaction, the slope relating a predicted dependent variable to the aptitude is different in each group. That was exactly the point developed earlier when it was

explained how aptitude-treatment interaction terms that were embedded in complicated regression equations could be interpreted by simplifying the original regression results. Here, this same idea of relations that change as one considers different groups is extended to terms in a regression equation, like an aptitude-notetaking-treatment interaction, where part of the interaction is a continuous variable rather than a code for a group. Now, not only does the slope relating a predicted dependent variable to an aptitude change depending on which group is being considered, but this slope also changes within the group of students as their scores change on the notetaking variable change. Although it is a long sentence to read, the statement that captures this idea is: When there is an aptitude-notetaking-treatment interaction, the relationship between students' aptitude and the predicted dependent variable changes as one considers different groups; and within a group of students, the relation between students' aptitude and the predicted dependent variable changes as one considers different scores on the notetaking variable. Thus, to fully appreciate the nature of an aptitude-notetaking-treatment interaction, more than one substitution needs to be made for the value of the notetaking variable so that changes in the relationship between the aptitude and the dependent variable can be displayed.

Before going on, a further question should be answered: Will there also be different y-intercepts for different values of a notetaking variable? It depends on the composite that forms the y-intercept. In the example being used here (see expression 10), the variable Prior Objective Achievement is not included in the composite that forms the y-intercept. So, in this case, the y-intercept will not change because there are no different values of Prior Objective Achievement to be substituted in the composite that forms the constant. However, it is possible for the composite that forms the y-intercept to include a variable that is also included in the second group of terms that produces the regression coefficient. When this is the case, there is a different y-intercept corresponding to each different value of the regression coefficient.

There are no well-accepted guidelines for choosing the number of different regression coefficients, and possibly their corresponding y-intercepts, that one needs to examine complex interactions fruitfully. In the analyses reported here, these complex interactions include aptitude-notetaking-treatment interactions or aptitude-aptitude-treatment interactions. Some researchers (e.g., Peterson, 1977) produce three-dimensional graphs of such interactions, thereby avoiding the issue of how many values to choose by displaying an infinite range of such coefficients. We find these figures useful, providing that the reader can interpret them. An alternative, illustrated by Cohen and Cohen (1975, section 8.4), is to produce regression equations for each group at three points along the scale of the "extra" interacting variable (the notetaking variables in the previous example): the mean minus one standard deviation, the mean, and the mean plus one standard deviation. We have followed this latter course in this report, judging it easier to present and to interpret. Our choice also seems more defensible in the light of small sample sizes for which an infinite range of coefficients might suggest a smooth change where such relationships may not be justified by our data. A disadvantage of our choice, however, is that standard deviations are sensitive to outliers, although the presence of outliers would distort three-dimensional graphs as well. Further, one should not discard outliers from a data set unless it can be determined that the outliers represent a different population than the scores constituting the remainder of the distribution. We assume here that outliers simply reflect the absence of intermediate values due to small samples. Thus, when we describe aptitude-aptitude-treatment interactions or aptitude-notetaking-treatment interactions, we present tables describing within-group regression slopes and intercepts at these three points of the scale for the interacting variable.

The final issue to be explained before turning to the results is why only one effect size statistic is reported per analysis. It may seem intuitive that an effect size describing differences between group means could be reported for every term in the regression analyses that includes the group code. For example, in

expression 9, it may appear reasonable that four effect sizes could be calculated, one for each of the terms labeled c, d, e, and f. The reason that this is not possible is given by the same principle for regrouping terms as has been demonstrated previously. When, in expression 9, terms labeled a and b are grouped to form a constant, and terms labeled c through f are collected to create a regression coefficient, there is only one equation to describe the predictor Group, namely:

(13) Predicted Total Objective Score =

$$[\text{Constant} + b_1 (\text{ICR Aptitude})] + [b_2 + b_3(\text{ICR Aptitude}) + b_4 (\text{Prior Objective Achievement}) + b_5(\text{Prior Objective Achievement} \times \text{Total Notes})]\,\text{Group}.$$

Substituting group means for the aptitude and notetaking variables, and a code for the Group produces only one predicted mean score per group. Hence, only one effect size can be computed comparing the two groups, calculated as:

(14) $E = (Y'_T - Y'_c) \,/\, MS_{residual}$

Regardless how the terms in expression 9 are arranged, for example as in expressions 10 plus 11, they can always be rearranged to give expression 13. In other words, when there are only two groups, there can be only one effect size when group means are substituted for continuous variables contained in the regression equation. This differs from the results of Study III, where the four vectors representing contrasts among five groups were all free to remain in any particular regression equation. In that case, one separate effect size could be calculated for each contrast. But when a particular contrast was also involved in an aptitude-treatment interaction, only one effect size was calculated for both the contrast main effect and the aptitude-treatment interaction. With these statistical explanations behind us, we turn now to the results.

Class A

Descriptive statistics for the dependent variables in Class A are reported in Table 18. Differential effects on students' achievement due to the treatment were observed for every dependent variable except Non-Tagged Objective Items. Although these effects were statistically reliable, the effect size statistics

TABLE 18

Means, Standard Deviations, and Correlations Among All Variables in Class A
Study IV

	Trained	Comparison	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. ICR Aptitude	8.90[1] 1.66	8.45 2.34	18[2]	54	37	33	42	14	16	06	27	31	11	46	48	17
2. Prior Objective Achievement	10.00 2.21	9.82 1.17		56	02	05	-06	-35	-37	-18	38	25	48	14	00	21
3. Prior Essay Achievement	13.40 3.72	12.09 4.30			34	33	33	14	13	12	49	55	24	37	06	49
4. Total Notes, Objective Items	37.50 14.23	33.55 15.05				99	91	44	44	28	17	22	03	59	32	54
5. Tagged Notes, Objective Items	24.90 9.80	22.91 11.48					83	42	42	26	14	18	01	61	41	49
6. Non-tagged Notes Objective Items	12.60 4.86	10.63 4.11						44	45	28	23	29	06	45	07	60
7. Total Notes, Essay Concepts	20.40 4.43	23.73 10.52							95	75	-20	-07	-36	30	20	21
8. Tagged Notes, Essay Concepts	14.10 3.18	17.73 7.99								50	-17	-06	-30	19	15	12
9. Non-tagged Notes, Essay Concepts	6.30 2.21	6.00 3.55									-20	-08	-35	43	24	39
10. Total Objective Scale	14.20 4.13	13.18 4.12										94	80	-08	-23	12
11. Tagged Objective Items	9.50 3.14	9.09 2.77											54	02	-18	21
12. Non-tagged Objective Items	4.70 1.49	4.09 1.87												-23	-24	10
13. Total Essay Scale	1.70 1.49	1.91 2.17													76	69
14. Tagged Essay Scale	.70 .67	1.18 1.72														06
15. Non-tagged Essay Scale	1.00 1.41	.73 1.01														

1 Upper numbers are means; lower numbers are standard deviations.

2 are omitted on correlations.

suggest that educationally important differences were not found (see Table 19). In particular, on the Total Objective Scale, students' achievement in the trained group corresponded only to the 53rd percentile relative to the average student in the comparison group. On the other dependent variables, differences in these relative terms were indistinguishable. However, these omnibus contrasts require qualification on the basis of the interactions that were common among these analyses. Table 20 shows the calculations for these interactions as described in the previous section.

The interaction involving Tagged Objective items as the dependent variable provides an example of the complex results of training students to use cognitive strategies in regular classroom lessons. Consider the interaction using notetaking (X_2 in Table 20) as a grouping variable rather than ICR aptitude. Either set of equations in Table 20 provides the same information and reflects the same statistical result reported in Table 19. The upper set where notetaking is the grouping variable provide a clearer depiction of the effects of training than the alternative. Note that at the means on notetaking for both groups, the differences of slope for the two groups are inconsequential (see Table 19). However, high notetakers with high ICR aptitude (at +1 standard deviation) in the trained group had high scores on Tagged Objective items, while their trained colleagues who had high notetaking scores but low ICR aptitude had low achievement scores. The opposite result occured in the comparison group. Here, high aptitude, high notetakers had poor achievement while high aptitude, low notetakers were more successful achievers.

These results can be interpreted as follows. The ICR aptitude test was designed to indicate the degree to which students were able to interpret their teacher's intentions for students' cognitive processing. The test was administered in early spring, after the students had been working with their teachers for about seven months. Assuming that the teachers had not drastically altered their teaching methods nor their intentions for student cognition after we had begun working with them, the ICR aptitude measure assessed an historically developed set of communications between

TABLE 19

Backward Selection Regression Analyses for Class A
Study IV

Dependent Variable	Predictors	β	F/t[1]	p	% variance[2]	E
Total Objective Scale	Total equation		6.04	.01	.34	.07
	Prior Objective Achievement	.33	1.81	.09	.15	
	ICR Aptitude x Total Notes x Group	.51	2.74	.01	.26	
	Constant	(13.53)	18.71	< .01		
Tagged Objective Items	Total equation		4.93	.04	.16	.01
	ICR Aptitude x Tagged Notes Group	.45	2.22	.04	.21	
	Constant	(9.24)	16.04	< .01		
Non Tagged Objective Items	Total equation		5.81	.03	.19	–
	Prior Objective Achievement	.48	2.41	.03	.23	
	Constant	(4.38)	13.25	< .01		
Total Essay Scale	Total equation		3.48	.08	.11	.00
	ICR Aptitude x Group	-.39	-1.87	.08	.15	
	Constant	(1.89)	4.98	< .01		
Tagged Essay Scale	Total equation		7.05	.02	.23	-.01
	ICR Aptitude x Group	-.52	-2.66	.02	.27	
	Constant	(1.02)	4.04	< .01		
Non-tagged Essay Scale	Total equation		5.65	.01	.32	.00
	Prior Objective Ach x Group	-.41	2.18	.04	.16	
	Non-tagged notes x Group	-.40	-2.12	.05	.15	
	Constant	(.86)	3.96	< .01		

[1] F-statistics are reported for the regression equation containing all predictors remaining in the equation ($p < .10$); t-statistics are reported for each predictor's slope coefficient.

[2] The figure reported for the Total equation is R^2 adjusted for shrinkage. Those associated with each predictor are squared part (or semipartial) correlations.

TABLE 20

Simplified Regression Equations for Interpreting Interactions from Analyses in Class A
Study IV

Dependent Variable	Predictor (X)	Within-group Regression Equations
Total Objective Scale	ICR Aptitude (X_1) x Total Notes (X_2) x Group	$E' = 13.60 - 1.04X_1$ at $-1sd\ X_2$ $E' = 13.60 + .18X_1$ at $M\ X_2$ $E' = 13.60 + 1.39X_1$ at $+1sd\ X_2$ $C' = 13.48 - 1.45X_1$ at $-1sd\ X_2$ $C' = 13.48 + .02X_1$ at $M\ X_2$ $C' = 13.48 + 1.24X_1$ at $+1sd\ X_2$ $E' = 13.60 + .16X_2$ at $-1sd\ X_1$ $E' = 13.60 + .02X_2$ at $M\ X_1$ $E' = 13.60 - .12X_2$ at $+1sd\ X_1$ $C' = 13.48 + .22X_2$ at $-1sd X_1$ $C' = 13.48 + .02X_2$ at $M\ X_1$ $C' = 13.48 - .18X_2$ at $+1sd\ X_1$
Tagged Objective Items	ICR Aptitude (X_1) x Tagged Notes (X_2) x Group	$E' = 9.24 - .67X$ at $-1sd\ X_2$ $E' = 9.24 + .08X$ at $M\ X_2$ $E' = 9.24 + .82X$ at $+1sd\ X_2$ $C' = 9.24 + .95X$ at $-1sd\ X_2$ $C' = 9.24 + .02X$ $M\ X_2$ $C' = 9.24 - .80X$ at $+1sd\ X_2$ $E' = 9.24 - .11X$ at $-1sd\ X_1$ $E' = 9.24 + .02X$ at $M\ X_1$ $E' = 9.24 + .36X$ at $+1sd\ X_1$ $C' = 9.24 + .19X$ at $-1sd\ X_1$ $C' = 9.24 + .02X$ at $M\ X_1$ $C' = 9.24 - .16X$ at $+1sd\ X_1$
Total Essay Scale	ICR Aptitude x Group	$E' = 1.89 - .36X$ $C' = 1.89 + .36X$
Tagged Essay Scale	ICR Aptitude x Group	$E' = 1.03 - .34X$ $C' = 1.03 + .34X$
Non-tagged Essay Scale	Prior Objective Achievement x Group	$E' = .83 + .29X$ $C' = .83 - .29X$
	Non-tagged Notes x Group	$E' = .88 - .16X$ $C' = .88 + .16X$

Note: For aptitude-aptitude-treatment interactions or aptitude-notetaking-treatment interactions, the regression equations at the mean of the variable for which group scores were substituted are equivalent to the aptitude-treatment interaction or the notetaking-treatment interaction from the preceding table in which the grouping variable does not appear.

Microfilmed From Best Available Copy

teachers and students. Presumably, the students who scored high in ICR aptitude were those who had understood these communications more successfully than those students who had low scores. But understanding these intentions is not the same as being able to execute the instructional response that teachers intend. The training may have provided this capability, but only if the training were used. Thus, students who were successful at perceiving these intentions, but were unable or unwilling to execute the response had relatively poor achievement. This was the high aptitude, low notetaking trained group. On the other hand, some of the high aptitude, high notetaking students who had been trained now were capable of using and willing to employ a rather sophisticated cognitive strategy to learn from teaching. Their high notetaking scores constitute _prime facie_ evidence that they used these strategies specifically in association with the teacher's signals to use them with content that was to be tested.

In the comparison group, high ability, high notetakers performed poorly on the achievement test. Recall that their notetaking did not necessarily represent the execution of an appropriate cognitive response. Rather, it simply indicated that these comparison group students had processed instruction and written something corresponding to tested content in their notebooks, but not necessarily at the point of the lesson when this information was tagged by the teacher. High scores indicated more test appropriate notetaking, but not necessarily the more frequent execution of an appropriate strategic cognitive response. Thus, these students' notes may be reflecting a high degree of sensitivity to any of a host of varied instructional stimuli that they perceived somewhere in the lessons, but the inappropriate or incomplete execution of a strategy for cognitively elaborating the information for storage in long term memory due to allocating too much cognitive resource to notetaking. As a result, they were less successful on the test. This may have happened despite the fact that their high ICR aptitude scores indicated that they could recognize situations in which appropriate cognitive responses were being signalled. On the other hand, high aptitude, low notetaking students in the comparison groups were perceiving the teacher's intentions for

cognitive processing, but were not distracted by the task of taking notes that were not necessarily representative of appropriate cognitive operations, as was notetaking in the trained group. These students, then, may have been attending to the important information to be tested later, and elaborating this information for storage in long term memory by using idiosyncratically developed cognitive responses about which we have no data.

The form of the ICR Aptitude x Total Notes score x treatment group interaction on the Total Objective achievement score is quite different than the one discussed previously. The difference in intercepts between the two groups is slight numerically, as well as practically given the small effect size (E = .07) at the joint mean of ICR aptitude and Total Notes. However, note the regression coefficients for each group that show the relationship of ICR Aptitude to predicted achievement scores at three points (-1 sd, M, +1 sd)along the Total Notes scale (X_2 for the analyses reported for the Total Objective scale in Table 20). What is strikingly apparent is that the regression slopes for the two groups are negligibly different both at the mean and one standard deviation above the mean (-.15 and -.16, respectively) for total notes taken. But at -1 standard deviation the difference was -.41. Herein lies the locus of the three-way interaction. High ICR aptitude students have higher total scores on the multiple choice test in this class if they take more notes, regardless of treatment group. For students who took an average number of notes, their achievement was correlated to aptitude. Again, this was true regardless of treatment group. But treatment did make a difference for high aptitude students who took relatively few notes. For these students, achievement was relatively low whether they were trained or not. However, the high aptitude, low notetaking trained students performed better than their untrained counterparts, indictating that training did influence the cognitive responses of these students.

Turning to the essay dependent variables, the interaction of ICR Aptitude and Group was nearly identical on the Total Essay Scale and the Tagged Essay Scale. In the trained groups, students' predicted achievement was negatively related to ICR Aptitude. The opposite was true of students in the comparison group.

On the Non-Tagged Essay Scale, two interactions were observed. Prior Objective Achievement was proportionally related to predicted essay scores in the trained group, and inversely related in the comparison group. In contrast, students in the trained group who took more rather than fewer notes corresponding to non-tagged information tested on the essay test were predicted to score lower on this essay scale. The reverse was the case in the comparison group where the amount of notes taken in non-tagged concepts was positively associated with predicted achievement.

Class B

Descriptive statistics for all variables are displayed in Table 21. In this class, achievement on the essay test and its two subscales was not affected differentially by the experiment since no predictors involving the contrast between groups accounted for statistically reliable variance in these three dependent variables ($p > .10$). There were, however, statistically reliable and educationally important relations among students' scores on the objective tests involving training and both prior achievement and notetaking variables (see Table 22). On all three objective achievement scales, trained students generally outscored comparison group students. Effect size statistics based on the regression equations show that an average trained student scored at the 92nd percentile on the total objective test and at the 93rd percentile on the non-tagged objective items compared to an average untrained counterpart. Although the same direction of effect was found on the tagged objective scale, the magnitude of this effect was considerably less. An average trained student scored only at the 56th percentile relative to an average comparison group student.

Table 23 displays the simplified regression equations for statistically reliable interaction terms that were retained in these backward selection regression analyses. For the two dependent variables showing large effect sizes (Total Objective Scale and Non-tagged Objective Items), opposite relationships between aptitudes and achievement resulted for the two groups of students. When the dependent variable was Total Objective Achievement, Prior Objective Achievement was inversely related to achievement in the trained

TABLE 21

Means, Standard Deviations, and Correlations Among all Variables in Class B, Study IV

	Trained	Comparison	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. ICR Aptitude	7.75[1] 1.58	6.85 1.77	12[2]	18	-14	-14	07	-24	-12	-30	21	17	19	-25	21	-17
2. Prior Objective Achievement	7.50 2.14	6.43 2.70		59	26	25	27	06	-09	09	57	51	47	39	51	13
3. Prior Essay Achievement	4.37 2.20	2.71 1.80			53	55	48	18	06	20	87	74	77	40	59	07
4. Total Notes, Objective Items	24.38 20.87	30.00 11.87				97	46	65	47	67	50	34	54	55	53	38
5. Tagged Notes, Objective Items	12.00 10.52	15.42 6.58					86	71	51	74	44	31	46	51	49	37
6. Non-tagged Notes, Objective Items	12.38 10.81	14.57 6.34						54	39	55	53	36	58	55	55	37
7. Total Notes, Essay Concepts	7.13 6.98	22.29 7.52							89	99	06	01	10	15	07	20
8. Tagged Notes, Essay Concepts	1.00 1.41	4.00 1.00								85	-05	-07	-02	-01	-08	06
9. Non-tagged Notes, Essay Concepts	6.13 5.87	18.29 6.90									08	02	13	19	10	22
10. Total Objective Scale	11.13 3.40	8.29 3.15										88	86	53	71	15
11. Tagged Objective Items	4.75 1.49	3.42 2.51											51	40	64	01
12. Non-tagged Objective Items	6.38 2.07	4.86 1.57												52	59	26
13. Total Essay Scale	5.75 3.01	4.43 2.57													85	82
14. Tagged Essay Scale	2.87 1.55	1.57 1.81														40
15. Non-tagged Essay Scale	2.88 1.73	2.86 1.57														

1 Upper numbers are means; lower numbers are standard deviations.

2 D[illegible]re omitted on correlations.

TABLE 22

Backward Selection Multiple Regression Analyses for Class B, Study IV

Dependent Variable	Predictors	B	F/t[1]	p	% variance[2]	E
Total Objective Scale	Total equation		27.07	<.01	.79	1.43
	Prior Essay Achievement	.83	6.70	<.01	.68	
	Prior Objective Achievement x Group	-.24	-1.96	.07	.06	
	Constant	(9.99)	23.45	<.01		
Tagged Objective Items	Total equation		4.89	.02	.53	.16
	Prior Objective Achievement x Group	-.58	-2.97	.01	.30	
	ICR Aptitude x Group	.43	2.04	.07	.14	
	Tagged Notes x Group	.43	2.21	.05	.16	
	ICR Aptitude x Tagged Notes x Group	-.60	-2.81	.02	.27	
	Constant	(4.29)	10.43	<.01		
Non-tagged Objective Items	Total equation		7.95	<.01	.71	1.45
	Prior Essay Achievement	.41	2.09	.07	.09	
	Non-tagged Notes	.57	2.78	.02	.16	
	Group	.32	1.87	.09	.07	
	Prior Essay Achievement x Group	.34	1.90	.09	.07	
	Non-tagged Notes x Group	-.45	-2.24	.05	.10	
	Constant	(5.23)	15.98	<.01		
Total Essay Scale	Total equation	--	--	--	--	--
Tagged Essay Scale	Total equation		6.78	.02	.29	--
	Prior Essay Achievement	.59	2.61	.02	.34	
	Constant	(2.27)	5.96	<.01		
Non-tagged Essay Scale		--	--	--	--	--

1 F-statistics are reported for the regression equation containing all predictors remaining in the equation ($p < .10$); t-statistics are reported for each predictor's slope coefficient.

2 The figure reported for the Total equation is R^2 adjusted for shrinkage. Those associated with each predictor are squared part (or semipartial) correlations.

TABLE 23

Simplified Regression Equations for Interpreting Interactions from Analyses in Class B, Study IV

Dependent Variable	Predictor (X)	Within-group Regression Equations
Total Objective Scale	Prior Objective Achievement x Group	$E' = 11.05 - .37X$ $C' = 8.79 + .37X$
Tagged Objective Items	Prior Objective Achievement x Group	$E' = 4.44 - .51X$ $C' = 4.24 + .51X$
	ICR Aptitude (X_1) x Tagged Notes (X_2) x Group	$E' = 2.77 + 2.19X_1$ at $-1sdX_2$ $E' = 3.86 + .77X_1$ at MX_2 $E' = 4.96 - .64X_1$ at $+1sdX_2$ $C' = 3.50 + .08X_1$ at $-1sdX_2$ $C' = 3.80 - .31X_1$ at MX_2 $C' = 4.87 - 1.69X_1$ at $+1sdX_2$
		$E' = 3.38 + .26X_2$ at $-1sdX_1$ $E' = 4.26 + .05X_2$ at MX_1 $E' = 5.14 - .16X_2$ at $+1sdX_1$ $C' = 5.25 - .41X_2$ at $-1sdX_1$ $C' = 4.25 - .17X_2$ at MX_1 $C' = 3.27 + .07X_2$ at $+1sdX_1$
Non-tagged Objective Items	Prior Essay Achievement x Group	$E' = 5.82 + .72X$ $C' = 4.89 + .03X$
	Non-tagged Notes x Group	$E' = 6.40 + .00X$ $C' = 4.59 + .23X$

Note: For aptitude-aptitude-treatment interactions of aptitude-notetaking-treatment interactions, the regression equations at the mean of the variable for which group scores were substituted are equivalent to the aptitude-treatment interaction or the notetaking-treatment interaction from the preceding table in which the grouping variable does not appear.

group, but positively related to achievement in the untrained group. Thus, on the objective items as a whole, training allowed low aptitude students to achieve more like high aptitude students who presumably used natural cognitive strategies during instruction. These two regression equations have their point of intersection at +1.28 standard deviation units or the 90th percentile on the Prior Objective Achievement Scale. We conclude on this basis that training does not hurt high aptitude students' performance.

Students' performance on the non-tagged information tested by objective items was positively related to their prior essay achievement in the trained group, but unrelated in the comparison group. In other words, students who demonstrated more aptitude for learning information relevant to essay items by displaying that knowledge on a prior essay test profited more from training than did students lower in this aptitude. However, the point at which the regression equations for the two groups intersect is at -.63 standard deviation units or the 24th percentile on the Prior Essay Achievement Scale. This suggests that it would not be very common for a low aptitude student who was trained to make intended cognitive responses to achieve less than an untrained student. Thus, training boosted students' ability to learn information tested by non-tagged objective items overall, and especially for students' with higher aptitude for learning essay-related information.

The regression equations describing the interaction of the non-tagged notes score with treatment revealed that students who were trained to make intended cognitive responses did not profit further on non-tagged information tested by objective items by taking more notes while untrained students did. However, taking into account the differences in intercepts for these equations shows that untrained students must exceed at least the 82 percentile in notetaking (+.90 standard deviation units) to match or better the achievement levels reached by trained students. It appears, then, that the majority of untrained students would profit from training because they could take fewer notes during class (about one-third as much) and still demonstrate relatively higher achievement.

Several relations need to be considered jointly in the interpretation of the interaction terms that emerged in the analysis of Tagged Objective Items. First, the interaction of students' prior objective achievement and treatment took much the same form as this same interaction that emerged in analysis of the Total Objective Scale, except that the point of intersection was at the 53rd percentile (+.08 standard deviation units on the Prior Objective Achievement Scale). Thus, this interaction describing learning of information that was tagged is much more disordinal than the former. The present intersection suggests that training students to make intended cognitive responses when cued by the teacher's instructional stimulus interferes with high aptitude students' learning tagged information, but promoted low aptitude students' achievement on this scale. Indeed, training made low aptitude students perform as well as high aptitude untrained students.

The simple regression equations produced from the complex interaction of ICR Aptitude, notes made about information tagged by the teacher's instructional stimulus, and the treatment condition show several things (see Table 23). First, the y-intercepts for the equations in which ICR Aptitude is the predictor (the first of the two blocks of six equations associated with ICR Aptitude (X_1) x Tagged Notes (X_2) x Group in Table 23) progressively increases as students' scores increase from -1sd to +1sd on the variable measuring notes taken about information their teacher tagged with an instructional stimulus. Since the ICR Aptitude scores were deviated about the grand mean in these analyses, this progression of values for the y-intercepts indicates that as average aptitude students took more notes about tagged concepts, their predicted achievement increases also. However, the interaction of ICR Aptitude and Tagged Notes indicates that this relationship varies with the level of students' ICR Aptitude.

In the experimental group, the relation between ICR Aptitude and predicted achievement on information tagged by instructional stimuli becomes less positive as students took more notes about tagged information. In fact, for students whose notes scores for

tagged information were +1 standard deviation from the mean, there was a negative relationship between ICR aptitude and achievement on Tagged Objective Items. This suggests that students at low and average levels of aptitude for recognizing when to make appropriate cognitive responses profited from training when they took more notes about tagged information, but that students high in this aptitude would be predicted to score higher as they take fewer notes. However, when trained students took ~~relatively more~~ notes (+1sd), only those students with ICR Aptitude scores that exceeded the 69th percentile for their group had predicted achievement ~~scores~~ that fell below those for trained students who took an average amount of notes about tagged information. This crossover point is at the 81st percentile of ICR Aptitude scores when comparing frequent notetakers (+1sd) with less frequent ones (-1sd).

Thus, training students to make intended cognitive responses during lessons is somewhat beneficial in general, elevating their learning of tagged information tested by objective items to the 56th percentile relative to average comparison students. These results could be interpreted as follows. Suppose that high ICR Aptitude students do not provide behavioral indications of the intended cognitive response we trained them to use because they don't use this mediation when the teacher tags information. Rather, they substitute a cognitive response that they presumably have developed by virtue of the fact that they exhibited high ICR Aptitude scores. Following these assumptions, it is possible that these students have rejected our version of cognitive processing in favor of processes they prefer and have developed on their own (see Winne & Marx, 1979). In contrast, lower and middle ICR Aptitude students who used our intended cognitive response when cued by their teacher (the +1sd equation of the first block of six equations) are predicted to score very well indeed on this information as indicated by the highest y-intercept (4.96) and a negative regression coefficient (-.64) for this equation.

Although training students to make intended cognitive responses was beneficial overall in this class, the complex interaction of ICR Aptitude, Tagged Notes, and Group reveals that the effects of

training are not always superior to no training. In fact, training is predicted to enhance achievement of information that the teacher tagged with an instructional stimulus only under certain conditions in this class. For students who took relatively fewer notes, no training was better if ICR Aptitude scores were below the 58th percentile. For those who took an average number of notes or a relatively larger quantity of notes about tagged information, no training was better for students who scored below the 49th percentile or the 48th percentile, respectively.

Class C

On objective measures of achievement for students in Class C, training in the use of intended cognitive responses was reliably related to learning. After scores had been residualized for prior achievement, students in the trained group outscored their untrained peers on the Total Objective Scale and on the Non-tagged Objective Items, reaching predicted averages of the 60th percentile on the former and the 76th percentile on the latter in comparison to predictions for an average student in the comparison group (see Table 24 for descriptive statistics and Table 25 for summaries of regression analyses). Thus, the training yielded a practical, as well as a statistically reliable effect. However, there were no statistically reliable differences identified between trained and untrained students on objective items that tested information that the teacher had tagged with instructional stimuli.

On the essay measures of learning, ICR Aptitude was a consistent variable that influenced relations among achievement, treatment, and in two of the three cases, notetaking by students. As well, the simple comparison of the two groups' predicted mean scores revealed that training deleteriously influenced essay-related learning overall and learning concepts that had been tagged by the teacher. In these two instances, average predicted scores for trained students fell to the 16th percentile on the Total Essay Scale and the 6th percentile on the Tagged Essay Scale, compared to the predicted average comparison student. The opposite effect occured on the Non-Tagged Essay Scale, however. On this scale, training aided learning, boosting predicted average achievement to

TABLE 24

Means, Standard Deviations, and Correlations Among all Variables in Class C, Study IV

	Trained	Comparison	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. ICR Aptitude	8.50[1] 1.72	9.83 .98	26[2]	31	33	29	31	-08	-11	-02	17	24	02	-14	13	-30
2. Prior Objective Achievement	8.30 2.58	8.83 1.83		53	43	22	55	42	41	33	79	68	75	58	61	44
3. Prior Essay Achievement	4.00 1.63	5.67 2.25			48	43	46	30	18	38	65	67	45	75	83	55
4. Total Notes, Objective Items	41.10 15.50	36.50 20.21				91	93	63	54	59	53	41	58	40	49	26
5. Tagged Notes, Objective Items	17.40 8.34	16.00 9.61					71	65	58	59	49	46	41	37	39	28
6. Non-tagged Notes, Objective Items	23.70 9.21	20.50 11.13						51	43	50	50	31	64	37	51	21
7. Total Notes, Essay Concepts	31.20 9.30	26.50 15.23							92	87	51	43	49	37	25	38
8. Tagged Notes, Essay Concepts	12.60 2.15	8.50 2.71								60	49	45	44	31	09	41
9. Non-tagged Notes Essay Concepts	18.60 3.41	18.00 8.90									40	31	44	35	39	26
10. Total Objective Scale	14.70 4.16	14.17 3.57										93	85	77	66	70
11. Tagged Objective Items	6.40 2.63	6.67 2.73											61	70	62	63
12. Non-tagged Objective Items	8.30 1.83	7.50 1.76												67	55	63
13. Total Essay Scale	3.70 3.23	4.17 3.82													84	92
14. Tagged Essay Scale	1.50 1.35	2.33 1.86														57
15. Non-tagged Essay Scale	2.20 2.30	1.83 2.23														

[1] Upper numbers are means; lower numbers are standard deviations.

[2] De[illegible] re omitted on correlations.

TABLE 25

Backward Selection Multiple Regression Analyses for Class C, Study IV

Dependent Variable	Predictors	β	F/t[1]	p	% variance[2]	E
Total Objective Scale	Total equation		14.72	<.01	.73	.26
	Prior Objective Achievement	.57	3.61	<.01	.23	
	Prior Essay Achievement	.48	2.80	.02	.14	
	Group	.34	2.27	.04	.09	
	Constant	(14.17)	26.80	<.01		
Tagged Objective Items	Total equation		9.69	<.01	.54	--
	Prior Objective Achievement	.5[illegible]	2.17	.05	.15	
	Prior Essay Achievement	.44	2.11	.06	.14	
	Constant	(6.50)	14.80	<.01		
Non-tagged Objective Items	Total equation		10.10	<.01	.61	.72
	Prior Objective Achievement	.79	4.87	<.01	.61	
	Group	.32	1.95	.07	.10	
	Constant	(7.86)	27.30	.01		
Total Essay Scale	Total equation		17.31	<.01	.78	-.99
	Prior Essay Achievement	-.40	4.83	<.01	.36	
	ICR Aptitude x Group	-.25	-1.95	.07	.06	
	ICR Aptitude x Total Notes x Group	.64	-3.00	.01	.14	
	Constant	(3.79)	10.29	<.01		
Tagged Essay Scale	Total equation		22.30	<.01	.81	-1.59
	Prior Essay Achievement	.77	6.58	<.01	.55	
	Tagged Notes	-.44	-2.79	.02	.10	
	ICR Aptitude x Tagged Notes x Group	-.56	-3.44	<.01	.15	
	Constant	(2.03)	11.20	<.01		
Non-tagged Essay Scale	Total equation		7.25	<.01	.56	.25
	Prior Essay Achievement	.75	3.96	<.01	.47	
	Group	.45	2.36	.04	.17	
	ICR Aptitude x Group	-.48	-2.74	.02	.22	
	Constant	(1.36)	3.23	.01		

[1] F-statistics are reported for the regression equation containing all predictors remaining in the equation ($p < .10$); t-statistics are reported for each predictor's slope coefficient.

[2] The figure reported for the Total equation is R^2 adjusted for shrinkage. Those associated with each predictor are squared part (or semipartial correlations.

the 60th percentile relative to an average untrained counterpart.

These simple findings are modified by interactions involving the essay test scores. Consider the Non-tagged Essay Scale first. On this dependent variable, trained students with average ICR Aptitude scores slightly outscored comparison students also having average ICR aptitude, as indicated by the difference in y-intercepts in Table 26. (Recall that ICR Aptitude has been deviated about its grand mean, thereby making the y-intercept correspond to average ICR Aptitude scores.) But the interaction of aptitude and treatment indicates that this description does not hold across the range of scores for ICR Aptitude. In fact, beyond the 79th percentile of ICR Aptitude scores, comparison group students are predicted to outscore their trained peers. Moreover, the negative regression coefficient relating ICR Aptitude and essay achievement of non-tagged concepts for the trained students (see Table 26) indicates the training put higher ICR Aptitude students at a disadvantage relative to low aptitude students with regard to achievement measured by the Non-tagged Essay Scale.

Analyses of the Tagged Essay Scale and the Total Essay Scale both revealed similar complex interactions involving ICR Aptitude, the relevant notetaking variable, and treatment. Consider the Tagged Essay Scale first. The y-intercepts for the simplified regression equations involving ICR Aptitude as the predictor (the third block of six equations in Table 26) show that predicted scores for comparison group students having average ICR Aptitude are at least as high or higher than those for trained students of average ICR Aptitude. Indeed, within any expected range of ICR Aptitude (+3sd) untrained students who take a mean number of notes about information tagged by instructional stimuli are always predicted on average to outscore trained students. Among students who took relatively fewer notes on these tagged concepts (-1sd), an average student trained to make intended cognitive responses whose ICR Aptitude scores exceeds the 65th percentile is predicted to score better than an average student in the comparison group. However, among students who took relatively more notes on the tagged concepts tested by essay items, an average trained student's ICR Aptitude needs to fall below the 30th percentile before the Tagged Essay

TABLE 26

Simplified Regression Equations for Interpreting Interactions from Analyses in Class C, Study IV

Dependent Variable	Predictor (X)	Within-group Regression Equations
Total Essay Scale	ICR Aptitude (X_1) x Total Notes (X_2) x Group	$E' = 3.12 + .25X_1$ at $-1sdX_2$
		$E' = 3.12 - .77X_1$ at MX_2
		$E' = 3.12 - 1.79X_1$ at $+1sdX_2$
		$C' = 4.90 - .98X_1$ at $-1sdX_2$
		$C' = 4.90 + .25X_1$ at MX_2
		$C' = 4.90 + 1.92X_1$ at $+1sdX_2$
		$E' = 3.20 + .02X_2$ at $-1sdX_1$
		$E' = 3.41 + .05X_2$ at MX_1
		$E' = 2.08 - .20X_2$ at $+1sdX_1$
		$C' = 3.62 - .24X_1$ at $-1sdX_1$
		$C' = 5.38 - .09X_1$ at MX_1
		$C' = 5.60 + .13X_1$ at $+1sdX_2$
Tagged Essay Scale	ICR Aptitude (X_1) x Tagged Notes (X_2) x Group	$E' = 2.13 + .52X_1$ at $-1sdX_2$
		$E' = 1.51 - .16X_1$ at MX_2
		$E' = .89 - .85X_1$ at $+1sdX_2$
		$C' = 3.68 - 1.14X_1$ at $-1sdX_2$
		$C' = 2.90 - .27X_1$ at MX_2
		$C' = 2.12 + .60X_1$ at $+1sdX_2$
		$E' = 1.66 + .05X_2$ at $-1sdX_2$
		$E' = 1.66 - .01X_2$ at MX_1
		$E' = 1.66 - .08X_2$ at $+1sdX_1$
		$C' = 2.65 - .04X_2$ at $-1sdX_1$
		$C' = 2.65 - .06X_2$ at MX_1
		$C' = 2.65 + .03X_2$ at $+1sdX_1$
Non-tagged Essay Scale	ICR Aptitude x Group	$E' = 1.84 - .72X$
		$C' = 1.23 + .72X$

Note: For aptitude-aptitude-treatment interactions or aptitude-notetaking-treatment interactions, the regression equations at the mean of the variable for which group scores were substituted are equivalent to the aptitude-treatment interaction or the notetaking-treatment interaction from the preceding table in which the grouping variable does not appear.

Scale score is predicted to exceed an average untrained students' achievement. Thus, students high in ICR Aptitude who operationalized training we gave them about making intended cognitive responses, insofar as performing the behavioral indicator was representative of these questions, were hurt. In contrast, students with higher ICR Aptitude who did less notetaking than we instructed were predicted to outscore untrained students.

Although the crossover points vary when Total Objective Essay scores are examined, the same relationships hold as were described for the Tagged Essay scores. Higher ICR Aptitude students who operationalized intended cognitive responses at either average or above average (+1sd) levels achieved lower scores relative to both untrained students in general, and to higher ICR Aptitude trained students who recorded relatively fewer notes (-1sd). Operationalizing training to make intended cognitive responses benefited only fairly low ICR Aptitude students in this class on the essay test.

Class D

In this class (see Table 27 for descriptive statistics), effects due to the treatment condition were observed for two of the three objective scales and one of the essay scales (see Table 28). On the Total Objective Scale, the interaction of Prior Objective Achievement with Group, and the intersection of Total Notes with Group complicate the simple interpretation than an average student in the trained group would be predicted to score at the 95th percentile relative to the averge of the comparison group distribution.

The interaction of prior achievement and treatment with the Total Objective Scale as the dependent variable is explored in Table 29. The within-group regression equations show that training lessened the dependence of the Total Objective Test scores on prior achievement (b=.47), compared to untrained students who used their natural cognitive processing strategies (b=1.48). However, predicted scores on the Total Objective Scale for trained students whose prior level of achievement exceeds the 82nd percentile begin to fall below those predicted for comparison group students. Thus, training students to make the cognitive responses

TABLE 27

Means, Standard Deviations, and Correlations Among All Variables in Class D, Study IV

	Trained	Comparison	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. ICR Aptitude	9.82[1] 2.86	7.75 2.43	45[2]	43	50	36	50	28	28	--	74	75	61	46	26	63
2. Prior Objective Achievement	8.91 2.66	7.38 3.02		75	44	38	41	34	34	--	61	48	67	52	44	42
3. Prior Essay Achievement	5.27 1.35	3.88 1.25			44	31	44	36	36	--	47	39	50	50	50	25
4. Total Notes, Objective Items	48.55 11.29	43.88 8.48				82	96	44	44	--	32	20	40	22	10	35
5. Tagged Notes, Objective Items	26.73 4.13	25.88 2.75					63	30	30	--	44	28	56	15	-04	47
6. Non-tagged Notes, Objective Items	21.82 7.95	18.00 6.99						44	44	--	22	14	28	23	15	26
7. Total Notes, Essay Concepts	45.00 4.45	41.63 5.66							1.00	--	32	31	29	39	39	20
8. Tagged Notes, Essay Concepts	36.00 4.45	32.63 5.66								--	32	31	29	39	39	20
9. Non-tagged Notes, Essay Concepts	9.00 0.00	9.00 0.00									--	--	--	--	--	--
10. Total Objective Scale	14.63 3.75	10.88 3.56										94	92	61	40	71
11. Tagged Objective Items	7.45 2.25	5.13 1.73											73	58	40	65
12. Non-tagged Objective Items	7.18 1.72	5.75 2.25												54	34	67
13. Total Essay Scale	3.91 2.21	3.38 2.00													92	65
14. Tagged Essay Scale	3.27 1.62	2.87 1.81														31
15. Non-tagged Essay Scale	.63 1.03	.50 .53														

1 Upper numbers are means; lower numbers are standard deviations.

2 are omitted on correlations.

TABLE 28

Backward Selection Multiple Regression Analyses for Class D, Study IV

Dependent Variable	Predictors	β	F/t[1]	p	% variance[2]	E
Total Objective Scale	Total equation		10.17	< .01	.75	1.87
	Prior Objective Achievement	.68	4.07	< .01	.23	
	ICR Aptitude	.50	3.07	.01	.13	
	Total Notes	-.59	-3.08	.01	.13	
	Group	.24	1.87	.09	.05	
	Prior Objective Achievement x Group	-.33	-2.16	.05	.06	
	Total Notes x Group	.60	3.21	.01	.14	
	Constant	(12.72)	25.41	.01		
Tagged Objective Items	Total equation		22.18	< .01	.54	--
	ICR Aptitude	.75	4.71	< .01	.57	
	Constant	(6.47)	17.96	< .01		
Non-tagged Objective	Total equation		12.83	< .01	.77	1.02
	Prior Objective Achievement	.80	5.48	< .01	.39	
	ICR Aptitude	.45	3.02	.01	.12	
	Non-tagged Notes	-.51	-3.22	.01	.13	
	Prior Objective Achievement x Group	-.43	-3.02	.01	.12	
	Non-tagged Notes x Group	.61	4.17	< .01	.23	
	Constant	(6.50)	27.15	< .01		
Total Essay Scale	Total equation		6.33	.02	.23	--
	Prior Objective Achievement	.52	2.52	.02	.27	
	Constant	(3.68)	8.78	< .01		
Tagged Essay Scale	Total equation		5.57	.03	.70	--
	Prior Essay Achievement	.50	2.36	.03	.25	
	Constant	(3.11)	9.11	< .01		
Non-tagged Essay Scale	Total equation		7.82	< .01	.43	.29
	ICR Aptitude	.53	2.88	.01	.26	
	ICR Aptitude x Group	.33	1.81	.09	.10	
	Constant	(.47)	3.01	.01		

[1] F-statistics are reported for the regression equation containing all predictors remaining in the equation ($p < .10$); t-statistics are reported for each predictor's slope coefficient.

[2] The figure reported for the Total equation is R^2 adjusted for shrinkage. Those associated with each predictor are squared part (or semipartial) correlations.

TABLE 29

Simplified Regression Equations for Interpreting Interactions from Analyses in Class D, Study IV

Dependent Variable	Predictor (X)	Within-group Regression Equations
Total Objective Scale	Prior Objective Achievement x Group	E' = 14.33 + .47X C' = 12.18 + 1.48X
	Total Notes x Group	E' = 14.62 + .01X C' = 9.59 - .48X
Non-tagged Objective Items	Prior Objective Achievement x Group	E' = 6.84 + .25X C' = 6.79 + .89X
	Non-tagged Notes x Group	E' = 6.95 + .03X C' = 5.32 - .31X
Non-tagged Essay	ICR Aptitude x Group	E' = .47 + .27X C' = .47 + .05X

was not beneficial for all; only students in the first four-fifths of the distribution of prior achievement are predicted to benefit from training.

In the case of the interaction between training and students' total notetaking during lessons, the within-group regression equations in Table 29 show that, for trained students, notetaking was essentially unrelated to their predicted achievement on the Total Objective Scale. In contrast, the more notes that comparison group students took, the lower their predicted achievement. Only comparison group students whose Total Notes score was below the 16th percentile would be predicted to outperform trained students. Thus, for all but a very few students, training in making intended cognitive responses during instruction was advantageous regardless of the extent to which those students operationalized their cognitive processing in the form of notes taken during lessons

The results for the Non-tagged Objective Scale parallel those for the Total Objective Scale. First, trained students outscored untrained students in general, averaging predicted scores at the 85th percentile of the untrained students' distribution. And, as before, the same types of interaction effects complicate this simple comparison. In the case of the interaction of prior achievement and treatment, training decreased the extent to which prior achievement predicted students' learning (see Table 29). However, the crossover point for these two within-group regression equations was much nearer the middle of the distribution of Prior Objective Achievement, namely, at the 53rd percentile. Thus, students below-average on prior achievement profited from training, but with respect to information tested on the objective scale that was not tagged by the teacher, it was more beneficial not to train students to make cognitive responses to teaching if their prior achievement was above average.

The interaction of students' notes about information not tagged by the teacher's instructional stimulus and treatment also was statistically reliable when the Non-tagged Objective Items served as the dependent variable (see Table 29). It showed that trained students would be predicted to outscore untrained students provided that they operationalized their cognition in the form of notes at

least above the 26th percentile. Below that level of notetaking about non-tagged information, trained students' predicted scores fell below their untrained peers.

Turning to the interaction of ICR Aptitude and Group identified in the analysis of the Non-tagged Essay Scale, the within-group regression equations (Table 29) show that above-average ICR Aptitude students profited from training, but trained students were predicted to score lower than comparison group students on this achievement scale if their ICR Aptitude was below average.

Class E

In this class, there were statistically and practically significant effects of training students to respond cognitively to instruction on all three of the objective measures of achievement. Descriptive statistics are displayed in Table 30. However, there were no statistically reliable effects of the treatment conditions on any of the essay measures of achievement. In the simplest terms, training was detrimental to learning, decreasing trained students' scores relative to comparison group students to the 14th percentile on the Total Objective Scale and the Non-tagged Objective items, and the 30th percentile on the Tagged Objective Items. However, in each case, these simple differences are complicated by aptitude-treatment interactions, notetaking-treatment interactions, or both. Table 31 reports the results of the backward selection regression analyses.

When the Total Objective Scale was the measure of achievement, the treatment manipulation interacted with students' prior objective achievement and with the extent to which they took notes about test-related information. In the case of prior objective achievement, the within-group regression equations (see Table 32) intersect at the 19th percentile on the scale for this aptitude. Thus, with the exception of all but a very few trained students wih low prior objective achievement, predicted achievement was higher for students in the control group than for students trained to make cognitive responses during teaching.

The opposite form interaction was found when students' notes was the interacting variable. However, because students' in the comparison group scored higher in general, only when trained students' Total Notes score exceeded the 92nd percentile was it

TABLE 30

Means, Standard Deviations, and Correlations Among All Variables in Class E, Study IV

	Trained	Comparison	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. ICR Aptitude	7.55[1] 2.07	8.07 2.84	21[2]	-03	33	23	34	08	-02	12	23	21	18	08	15	-04
2. Prior Objective Achievement	7.09 2.70	8.00 3.46		37	48	44	39	44	33	25	70	58	61	62	42	70
3. Prior Essay Achievement	4.27 2.10	3.43 2.10			15	07	18	33	11	34	50	44	39	55	49	48
4. Total Notes, Objective Items	43.90 8.51	40.93 9.68				85	87	58	38	38	38	41	22	40	45	24
5. Tagged Notes, Objective Items	19.36 4.90	18.21 5.58					50	54	31	40	39	38	27	26	24	22
6. Non-tagged Notes, Objective Items	24.55 5.66	22.71 5.09						46	35	26	27	33	11	44	53	19
7. Total Notes, Essay Concepts	37.36 6.20	36.43 7.05							68	65	45	45	30	36	22	43
8. Tagged Notes, Essay Concepts	13.91 5.89	15.86 4.31								-12	20	08	27	28	21	29
9. Non-tagged Notes, Essay Concepts	23.45 2.94	20.57 5.75									40	52	13	19	08	29
10. Total Objective Scale	9.81 2.52	12.50 4.62										86	82	60	42	66
11. Tagged Objective Items	4.91 2.34	5.93 2.64											42	53	43	52
12. Non-tagged Objective Items	4.91 1.64	6.57 2.41												47	28	59
13. Total Essay Scale	2.64 2.01	3.07 2.46													91	84
14. Tagged Essay Scale	1.00 1.41	1.14 1.51														53
15. Non-tagged Essay Scale	1.64 .92	1.93 1.27														

[1] Upper numbers are means; lower numbers are standard deviations.

[2] Decimals are omitted on correlations.

TABLE 31

Backward Selection Multiple Regression Analyses for Class E, Study IV

Dependent Variable	Predictors	β	F/t[1]	p	% variance[2]	E
Total Objective Scale	Total equation		8.32	<.01	.60	-1.06
	Prior Essay Achievement	.46	3.26	<.01	.17	
	Total Notes	.50	3.56	<.01	.21	
	Group	-.52	-3.97	<.01	.26	
	Prior Objective Achievement x Group	-.38	-2.32	.03	.09	
	Total Notes x Group	.29	1.82	.08	.05	
	Constant	(10.66)	19.75	<.01		
Tagged Objective Items	Total equation		5.27	<.01	.42	-.53
	Prior Essay Achievement	.55	3.38	<.01	.28	
	Tagged Notes	.45	2.81	.01	.19	
	Group	-.36	-2.27	.03	.13	
	ICR Aptitude x Group	.30	1.80	.09	.08	
	Constant	(5.45)	13.97	<.01		
Non-tagged Objective Items	Total equation		6.26	<.01	.52	-1.08
	Prior Essay Achievement	.30	1.91	.07	.07	
	Non-tagged Notes	.31	1.98	.06	.08	
	Group	-.50	-3.46	<.01	.06	
	Prior Objective Achievement x Group	-.59	-3.40	<.01	.23	
	Non-tagged Notes x Group	.27	1.78	.09	.06	
	Constant	(5.40)	16.55	<.01		
Total Essay Scale	Total equation		11.30	<.01	.46	--
	Prior Essay Achievement	.37	2.29	.03	.12	
	Prior Objective Achievement	.49	3.00	.01	.20	
	Constant	(2.88)	8.76	<.01		
Tagged Essay Scale	Total equation		7.15	.01	.20	--
	Prior Essay Achievement	.49	2.67	.01	.24	
	Constant	(1.08)	4.20	<.01		
Non-tagged Essay Scale	Total equation		22.62	<.01	.47	--
	Prior Objective Achievement	.70	4.76	<.01	.50	
	Constant	(1.80)	11.01	<.01		

[1] F-statistics are reported for the regression equation containing all predictors remaining in the equation (p <.10); t-statistics are reported for each predictor's slope coefficient.

[2] The figure reported for the Total equation is R^2 adjusted for shrinkage. Those associated with each predictor are squared part (or semipartial) correlations.

TABLE 32

Simplified Regression Equations for Interpreting Interactions from Analyses in Class E, Study IV

Dependent Variable	Predictor (X)	Within-group Regression Equations
Total Objective Scale	Prior Objective Achievement x Group	E' = 9.57 - .49X C' = 12.30 + .49X
	Total Notes x Group	E' = 9.23 + .35X C' = 12.62 + .09X
Tagged Objective Items	ICR Aptitude x Group	E' = 5.00 + .30X C' = 6.00 - .30X
Non-tagged Objective Items	Prior Objective Achievement x Group	E' = 4.69 - .42X C' = 6.40 + .42X
	Non-tagged Notes x Group	E' = 4.66 + .25X C' = 6.56 - .02X

predicted that training facilitated learning relative to the control experience. Thus, for most of the students in this class, training served to depress scores on the Total Objective Scale.

On the Tagged Objective Items, the within-group regression equations show that, for students in the comparison group, ICR Aptitude scores were inversely related to achievement on information tagged by instructional stimuli. But, for students who were trained to make cognitive responses, ICR Aptitude was proportionally related to achievement. Because the regression equations intersect at the 75th percentile for scores on the Tagged Notes Scale, only when trained students greatly manifested their cognitive processing of tagged information by use of the behavioral indicator, did their predicted achievement exceed the comparison group students' achievement.

The simplified regression equations involving either Prior Objective Achievement or Non-tagged Notes when Non-tagged Objective Items measured learning are nearly identical to those for the total Objective Scale (see Table 32). On this dependent variable, when trained students' prior achievement fell below the 26th percentile was their score predicted to exceed their average untrained counterpart's score. Also, only when trained students' non-tagged notetaking scores exceeded the 94th percentile was their achievement predicted to be higher than untrained counterparts.

CONCLUSIONS

As in Studies II and III and in earlier work (Winne, 1982b; Winne & Marx, 1979), we showed here that students could be taught to execute cognitive responses intended by the teacher. Further, the results of the work in the five classes studied here show that these cognitive responses could be employed, at least in large part, while the students were working on standard recitation teaching tasks in regular classrooms. This is encouraging, particularly because this cognitive strategy was so complex. However, the contrasts between the total notes scores and the behavioral indicator scores suggests that the trained students employed the cognitive response somewhat inconsistently throughout the lessons, at other times taking notes in a manner indicating that the intended cognitive response was not

used. Indeed, in Table 16, the average behavioral indicator scores corresponding to the objective test items ranged across the classes from 52 to 89 percent of the average Total Notes scores. For the essay tests, these figures were 44 and 79 percent. Presumably, at those times when the students' notes did not correspond to the behavioral indicator, they were using cognitive responses that they had developed prior to the training.

As in the previous studies, the relationship between the students' use of the cognitive response and achievement was inconsistent. While on some achievement scales and in some classes this relationship was strong (see Table 17), in other classes and with different measures of achievement, it was weak. For the trained students, the relationship was moderately strong (median r = .48) between use of the cognitive response and achievement on the objective tests, but not the essay tests (median r = .23). The opposite was true for the relationship between notetaking and achievement for the comparison students (see p. 185 and Table 17). However, the interpretation of the simple relationships between notetaking and achievement must be tempered by the presence of several notetaking x treatment interactions that emerged in classes B, D and E with the objective tests serving as the dependent variables. These are discussed below.

A summary of the various regression analyses performed in data in the five classes is reported in Table 33. Of the three aptitude measures, the two measures of prior achievement most often appeared as statistically reliable terms in the analyses. Interestingly while the measure of prior achievement on the essay test appeared more frequently in analyses in which the various essay measures constituted the dependent variable, both the objective and essay measures of prior achievement appear equally often when the objective measures served as dependent variables. ICR aptitude appeared as a predictor of achievement only in class D.

Notetaking appeared as a reliable predictor of at least one achievement measure in classes B, C, D, and E. The relationship was positive in classes B and C, but negative in classes D and E. These relationships, however, are moderated by the presence of several notetaking x treatment interactions, as referred to above.

TABLE 33

Summary of Statistically Reliable Terms from Backward Selection Multiple Regression Analyses for All Five Classes, Study IV

	APTITUDE[1]					INTERACTIONS						
	ICR (1)	Prior Object. Achiev. (2)	Prior Essay Achiev. (3)	Note-taking[1] (4)	Treatment Contrast[1] (5)	1 x 5	2 x 5	3 x 5	4 x 5	1 x 4 x 5	2 x 4 x 5	3 x 4 x 5
TOTAL OBJECTIVE SCALE	D+	A+ C+ D+	B+ C+ E+	D- E+	C+ D+ E-		B D E		D E	A		
TAGGED OBJECTIVE ITEMS	D+	C+	C+ E+	E+	E-	B E	B		B	A B		
NON-TAGGED OBJECTIVE ITEMS	D+	A+ C+ D+	B+ E+	B+ D- E+	B+ C+ E-		D E	B	B D E			
TOTAL ESSAY SCALE		D+ E+	C- E+			A C				C		
TAGGED ESSAY SCALE			B+ C+ D+ E+	C-		A				C		
NON-TAGGED ESSAY SCALE	D+	E+	C+		C+	C D	A		A			

Note: Letters indicate classes.

[1] The sign indicates the slope of the regression weight. For the treatment contrast, a positive slope indicates that the treatment group mean was reliably higher than the comparison group mean; a negative sign indicates a reliably higher mean for the comparison group.

The simple treatment contrast proved to be statistically reliable in four of the five classes for eight different regression analyses, seven of which involved one of the objective achievement scales. In class E, the contrast favored the comparison group. All other statistically reliable treatment contrasts favored the trained groups.

The most complex findings concern the aptitude x treatment and the aptitude x notetaking x treatment interactions. These are summarized in the seven columns on the right hand side of Table 33.

While ICR aptitude entered the regression analyses as a simple term only in class D, it appeared in either two- or three-way interction terms in all classes. From these data, it appears that the ICR aptitude measure has potential utility in uncovering relations in instructional research. However, like many aptitudes employed in the study of aptitude-treatment interactions, there appears to be a lack of reproducibility. For example, in the seven regression equations in which a two-way interaction appeared involving ICR aptitude, three indicated that high aptitude trained students and low aptitude comparison students scored higher, while four were of the opposite form.

However, some regularity appears in these results when the analyses are clustered by type of dependent variable. For the objective measures, only analyses involving the Tagged Objective Items included notetaking x treatment interactions (for classes B and E, see Table 33). Both of these analyses were of the same form, where trained students with high aptitude outperformed their untrained peers of equivilent aptitude, while untrained students with low aptitude outperformed their trained peers. Of the five times that notetaking x treatment interactions appeared in regression equations that had an essay measure as the dependent variable, four of the five had positive slopes for the comparison group and negative slopes for the treatment group. Generally, this showed that high aptitude comparison group students outscored high aptitude trained students. The exception was class D, where the high aptitude trained students were superior.

The prior achievement measures appeared in eight interaction terms with the treatment contrast. In seven of these regression equations an objective scale served as the dependent variable, with six involving the prior objective achievement items. In all of the equations where an objective scale served as the dependent variable and prior objective achievement interacted with treatment, the comparison group had a positive slope. The slopes for the trained groups in these classes (B, D, and E) were both positive and negative. In all cases, however, comparison group students with high prior objective achievement outperformed their trained peers, while trained students with lower prior objective achievement had higher predicted objective achievement scores than did untrained students. Thus, training served to attenuate the relationship between prior knowledge and achievement. This was particularly helpful to low aptitude students, for whom training served to increase scores to the level of untrained high aptitude learners.

On the one prior objective achievement x treatment interaction with an essay test as the dependent variable (non-tagged items), the high aptitude trained students and the low aptitude comparison group students scored highest. Thus, there was no evidence that training was able to help students overcome ability limitations on the essay test. Because this interaction occured in only one class and did not replicate, generalizations based on it is hazzardous. This is also true of the one interaction involving prior essay achievement and treatment.

Neither of the prior achievement measures appeared in a three-way interaction. These interactions all included the ICR aptitude test as the measure of aptitude. However, because of the strong likelihood that these effects are unstable, we judge it unwise to speculate more than we already have about these relations, especially in the light of the absence of reproducible effects across the five classes (see Table 33). Earlier discussions of these relations were included for speculative and exploratory reasons only. But, in accord with the caveat proffered in chapter 3, we choose not to lend further weight to these effects by summarizing them. That they were statistically reliable in the

individual analyses does not overcome our skepticism about their meaning. Though this same argument can be carried back to simpler effects, we have chosen not to employ it there because of the semblence of reproducibility that emerged for these findings. We point again to the caveat made earlier as sufficient warning about the statistical and the subsequent substantive robustness of the findings in toto.

CHAPTER 6

Conclusions and Implications

CHAPTER 6

Conclusions and Implications

At the end of each of the previous chapters, we have provided summaries or conclusions relating to the data presented in those chapters. Here, we draw together those findings and discuss conclusions and implications that arise from the project as a whole.

CONCLUSIONS

Research on teaching through the 1970s was based on a set of presuppositions about how teaching events were related to students' achievement. Along with methodological rules, these presuppositions constituted the process-product paradigm. One of these paradigmatic presuppositions underlying process-product research was that process variables, in the form of teacher behaviors such as statements of praise or higher-cognitive questions, directed students' on-the-spot cognitive operations on the content to be learned. Another presupposition was that the immediate products of these cognitive operations could be measured later by administering an achievement test to students. Thus process-product research sought to identify and explain relations between process variables operationally defined as teacher behaviors and students' achievement as represented by paper and pencil tests. It is important to note that, while researchers frequently used language that implied the existence of the intervening student cognitions relating these phenomena, rarely was the link between teacher behavior and student cognition explicitly discussed.

To interpret relations between these two categories of data, however, several other tacit assumptions had to supplement those already made. These tacit assumptions included at least the following: (1) Every occurrence of a uniquely operationally defined teacher behavior is associated with only one set of students' cognitive operations. (2) Every occurrence of any

operationally defined teacher behavior invariably results in a cognitive product that, once students have created it, could be reproduced invariantly on an achievement test. (3) Every student in a single classroom or treatment group experiences the same teacher behavior that the researcher operationally defined as the independent variable. This happens every time the teacher behavior occurs. These three tacit assumptions permit a researcher to perform logical and quantitative operations necessary to do things such as compute a correlation coefficient between a teacher's use of a teacher behavior and the class (or group) mean on a measure of achievement and interpret it in terms of a theory of learning (Gage, 1964) or motivation (see Winne, 1983a, b, for details).

A notable characteristic of each of these three tacit assumptions is that they make statements about some aspect of a student's covert cognitive operations in relation to observable events of instruction. As we argued in the first chapter, a series of empirical studies and logical analyses of the process-product paradigm produced some doubt about the paradigm's appropriateness. In particular, over the last decade a growing number of studies questioned the validity of each of the three tacit assumptions just mentioned. The studies comprising this project demonstrate clearly that such doubts are warranted. In an attempt to accomodate these findings, we adopted a new paradigmatic stance.

Whereas the process-product paradigm relied on assumptions about students' cognitive responses to instructional stimuli as defined a priori by a non-participant in instruction (i.e., the researcher), the cognitive mediational paradigm takes a more descriptive and more rigorous stand on this issue. Its rigor derives from the fact that it requires operational documentation about both students' cognitive operations during the course of instruction and the cognitive products that these operations yield. In this project, documentation about these phenomena was secured by two procedures: stimulated recall interviews of students in Study I, and the methodology of behavioral indicators in Studies II - IV.

As to the increased descriptive power of research guided by the cognitive mediational paradigm, this was achieved by explicitly accomodating the fact that students' vary in their naturalistic use of cognitive operations as well as in their achievement of cognitive products during instruction. Such accomodation took three forms in Studies II - IV of the project. The first form was to train students to perceive and act on instructional stimuli. This sharpened process-product methodology by reducing the range of variation in students' cognitive mediation that was observed in Study I. By training students in this manner, we attempted to control the goals students approached during instruction and the cognitive products they attained. This control was operationalized by determining a priori the cognitive operations students would draw on and by providing criteria in the form of behavioral indicators for them to judge their success at these tasks. As well, by tying these cognitive events to instructional stimuli delivered at specific points in the development of content in lessons, we attempted to control the amount of information students processed. Finally, by previously acquainting students with the exact structure and function of cognitive responses to chosen instructional stimuli, we tried to govern the immediacy of their cognitive mediations of these instructional stimuli.

The second form of accomodation we made to naturalistic variation in students' cognitive mediation of instruction concerned the assessment of cognitions at the time they occured during instruction. We did not assume that students necessarily would cognitively mediate instructional stimuli as had been presumed a priori in process-product research. Instead, we assessed the extent to which students actually used particular cognitive responses by collecting behavioral traces of their cognitive mediations. This measurement provided an index upon which to base interpretations about how instruction influenced immediate cognitive responses, which in turn related to achievement. This description served to reduce further some sources of variance that were extraneous to the propositions tested in these studies.

The third form of accomodation pertained to students' motivation. In the training sessions, special emphasis was made of the ways that the cognitive mediations being trained could support students when they responded to achievement test items. We assumed that demonstrating this congruence would motivate students to carry out nearly equivalent cognitive mediations in which they were trained, both during instruction and later when they responded to test items calling for those same cognitive mediations.

In gauging the qualities of these accomodations, students' behavioral indicators served as descriptive data for evaluating the extent to which the research in Studies II - IV accomodated the observations made in Study I. From another perspective, the latter three studies of this project tested rather than assumed the four components of successful cognitive mediation of instructional stimuli that Winne (1982a) identified. This provides both logical and methodological advantages to cognitive mediational research on teaching relative to process-product research.

In an effort to enhance or at least address teaching effectiveness in the project, we designed the cognitive mediations students were trained to use in Studies II - IV by building on prior findings from psychological research (e.g., Gagne, 1978) and from previous interpretations of process-product research on teaching. The hypotheses we advanced obviously evidenced our "informed commitment" (cf. Dunkin & Biddle, 1974, concerning less informed commitment) that students who were trained to carry out cognitive operations we identified should have learned more. Although this simple relation was only partially supported, and usually was complicated by interaction effects, there was some evidence that we were able to elevate trained students' achievement relative to untrained peers. Given the novelty of this approach to research, and the problems we encountered, these findings can be taken as indications about the potential of the cognitive mediational paradigm to foster research that is both theoretically and practically useful. Specifically, theoretical propositions about how students' cognitive mediations of

instruction influence learning can be examined by cognitive mediational research because these cognitive operations are defined and observed in the immediate environment of teaching rather than at a later time and indirectly by students' total scores on achievement tests (Snow, 1968). Practical utility also may derive from more exacting descriptions regarding what students can do cognitively when they participate in instruction, thereby prescribing how time on task can be filled to promote achievement (cf. Gage, 1978). But research that achieves these gains will need to solve several problems encountered in or raised by the studies reported here.

Problems in Cognitive Mediational Research

In many respects, a major focus of this project has been methodological. A broad sense of the term methodology is intended by this statement, meaning "the description, the explanation and the justification . . . of methods, and not merely the methods themselves" (Kaplan, 1964, p. 18). But a series of difficulties or weaknesses also characterized the methodology of this project. In this section, we discuss concerns of this nature in order to place appropriate bounds on the validity of the current findings and to note issues that future research will have to address.

Identifying cognitive mediations. Threats to the validity of conclusions from Study I have been summarized and debated by Nisbett and Wilson (1977) and Ericsson and Simon (1980). Our methodology fails to meet the standards of validity mentioned in these works, but neither is it so messy as to warrant outright dismissal (e.g., see Peterson, Swing, Braverman, & Buss, 1982 for one adoption of a nearly identical procedure). It can be considered as one source of potentially useful and valid descriptions of students' cognitive mediations of instruction, albeit one subject to some difficulties. Several of these difficulties warrant explicit attention.

Tobias (1982) and Calfee (1981) distinguish microprocesses and macroprocesses. The former are elementary cognitive operations. They may be under conscious control, such as rehearsal in the form of simple repetition, or they may not be

consciously regulated, such as discriminating vertical from angled components of letter forms. Macroprocesses, on the other hand, are assemblages of microprocesses that probably correspond to plans for dealing with instructional events. Reviewing information to organize it more compactly and to monitor the degree to which parts of this organization make sense is a likely candidate for an instructionally cued macroprocess.

Analyzing students' stimulated recall protocols to identify cognitive mediations is problematic in several respects. First, students are naive with respect to reflecting metacognitively, making their descriptions ambiguous in many, if not most instances. Second, when microprocesses and macroprocesses are highly automatic, students' are likely not to be aware of them and thus are unable to report their occurence. Third, to the extent that cognitive mediations involve macroprocesses that include orienting or cognitive responses that are to take place at a later time in the lesson, students may be unable to describe them.

Two approaches to dealing with these and related issues (Ericsson & Simon, 1980) can be proposed. One is to provide students with a brief education so that they can identify and describe cognitive events in terms commensurate with those used by researchers to describe cognition. This may help to lessen the gap in communication between students and researchers by lessening students' naivete. A second approach may be to devise instructional tasks that segment macroprocesses into either smaller macroprocesses or microprocesses that are of interest to the researcher, and requiring students to create behavioral traces that correspond to cognitive operations and cognitive products that can be examined (e.g., see Brown & Burton, 1978).

Of course, a third source that can be drawn on to identify microprocesses and macroprocesses that may characterize students' cognitive mediations of instruction is the vast literature of cognitive psychology. Tapping this source may need to be done with some caution, however, since this literature may exhibit some problems when identifying cognitive mediating operations. One caution, often given lip service but perhaps too rarely heeded in practice, is to consider possible developmental differences in

forms for macroprocesses. At the least, variations in the capacity of working memory may need to be addressed.

A second caution pertains to the use of nomothetic research strategies to study cognition. In most of these studies, data have been aggregated over multiple occurrences of a treatment variable (e.g., underlining terms in a text) and over students. Idiosyncracies in the cognitive activities of individuals as they respond to the same variable over time are aggregated into the uninformative category of residual variations. While such studies may describe "average" cognitive operations validly, it may be the case that the idiosyncratic variations would be a better source for identifying instructionally relevant cognitive mediations than models of the cognitive activities of a hypothetically average student, who, in fact, might not exist (see Winne, 1983a).

Tracing cognitive mediations. The behavioral indicators we designed to reflect students' cognitive mediations or some valid substitute for them are an essential component of cognitive mediational research on teaching. Several problems with obtaining these traces of students' cognitions can be described.

One difficulty with the behavioral indicator methodology regards fitting it into teaching so that it does not disrupt instruction, but supports it. There are several aspects to this problem. One of them is that, while students are engaged in producing a trace of cognitive mediations, they probably have to disengage cognitively from other potentially relevant instructional activities, e.g., listening to the teacher. In Studies II and III, it was relatively easy for us to create instruction that employed long teacher wait time. In Study IV, however, where instruction was more interactive than it was when it was delivered by videotape, as in Studies II and III, the teachers found it difficult to provide this wait time consistently following delivery of the instructional stimulus. This may have been an artifact of our design of Study IV. Had all rather than just half of the teacher's students been trained to engage specific cognitive mediations and produce behavioral indicators, we may havebeen able to avoid the need for teachers to keep non-trained students engaged. Similarly, they would not have had to

respond to the initiations of non-trained students during the wait time in which trained students were recording behavioral indicators. When instruction and students participating in that instruction are matched to take account of the wait time teachers must provide for gathering data in the form of behavioral indicators, this artifact may disappear.

A second aspect of lack of fit relates to students' propensity to produce behavioral indicators. At several points in our earlier discussions, we speculated that some students trained to use cognitive mediations may have done so but subsequently not provided a trace of those activities. Unless behavioral indicators per se have utility for students during the course of classroom learning, obtaining these data may be problematic. The motivational link we made between the cognitive mediating strategies students were trained to use during instruction and then again upon encountering a test item that could be answered using the same strategy, apparently did not induce the participants to record behavioral indicators at nearly the frequency that we would like in order to make stronger inferences regarding cognitive mediation of teaching. This problem needs attention in future work. Perhaps merely making these traces an automatic response, by considerable practice, will help to insure their production

A different problem with behavioral indicator data that was partly addressed in Studies II and III, but inadequately addressed in Study IV, are the links across specific occurrences of instructional stimuli, behavioral indicators, and achievement test items. Ideally, each triplet -- instructional stimulus, behavioral indicator, and one or a set of test items solely dependent on a cognitive product resulting from the execution of the intended cognitive response -- would be available for examination by the researcher. These kind of data, however, were not available in Study IV for two reasons. First, we could not invent a workable procedure by which students trained to provide behavioral indicators could log which occurrence of their teacher's instructional stimulus triggered their use of the cognitive mediational strategy. Thus, the link between the first

pair of items in the ideal triplet could not be assessed accurately for every student. Second, the generalized intended cognitive response designed for Study IV was representative of each classroom's milieu. But achieving this degree of representativeness and allowing freedom to the teachers to deliver all other aspects of instruction in their own way prohibited making the second link in the ideal triplet. This trade-off between internal validity and representativeness (Snow, 1974) lessens that informativeness of our data about the relations among items in the ideal triplet. Work to improve on our procedures in the context of classroom teaching is needed.

The last problem with behavioral indicator data is the most fundamental. If behavioral indicators of students' cognitive mediation of instructional stimuli are to be useful, they must have construct validity. That is, behavioral indicators must reflect the cognitive operations and cognitive products they are intended to reflect. In Studies II and III, the behavioral indicators we designed approximated the criteria listed on pp. 164-166. But Study IV provides an illustration of an experimental setting where other, more substantial problems can arise with the behavioral indicator methodology. The first of these derives from the fact that the behavioral indicator in Study IV was reactive both to a student's state of prior knowledge and to sequential events in the classroom. In part, this reactivity reflected the representativeness of the cognitive mediation that was designed based on data from the first phase of Study IV. But the trade off is that not all aspects of the behavioral indicator can be interpreted unambiguously. For example, if a student's behavioral indicator for Stage 1 is incomplete, it may be so because the instructional stimulus was not perceived, prior knowledge was lacking, or one or two parts of the cognitive mediations (see Figure 3) were not performed.

A related problem arises when behavioral indicators reflect macroprocesses that are not invariant, such as having several branches in a procedure. If, for example, a students' partially correct answer to a teacher's question resulted in the teacher going off on a tangent and never returning to the topic of the

question, students who were listening to this exchange would have needed to alter the second and third stage of their intended cognitive mediations and, hence, behavioral indicators. However, we did not provide a means for them to note this or other deviations from our one path through instructional events. Misinterpretations of incomplete behavioral indicators would result were this hypothetical situation to have occured in the study. Although better data collection procedures concerning all three items in the ideal triplet would help to counter this weakness, problems of this general type may be inherent in research of this genre unless the teacher follows a script very closely. Work in this vein merits attention in the near future.

Data analysis. Two problems are associated with the models of statistical analysis that were used in this research. Notable among these is the fact that, if students are discarded from a randomly constituted treatment group because their behavioral indicator data sustains the interpretation that they were not engaging in cognitive mediations as intended, the remaining group is non-random. This violates any traditional statistical procedure that rests partially on the assumption of random assignment (and retention). Both the internal validity of the study's design and the lack of bias in statistical estimates of parameters are threatened by systematic mortality of participants in a study.

A second difficulty, especially serious in Study IV, is the fact that every single classroom must be treated as a unique population because unmanipulated instructional stimuli and associated cognitive mediations vary across classes. This severely limits sample sizes with resulting losses to the robustness and precision of point estimate statistics such as means and partial regression coefficients.

These two problems, particularly the second, are associated with research on teaching derived from the process-product paradigm as well as research based on the cognitive mediational paradigm. As discussed earlier, teaching influences students in many ways. It is unlikely that all students in a classroom react cognitively in the same way to a particular teaching event.

Obviously, if this were the case, these students would not be reacting similarly to the range of teaching events that constitute a more global treatment. If the focus of one's theory of teaching is the reactions of individual students, then aggregating data across students reduces precision with regard to those individual reactions

These two problems underlie our strong position presented in Study II, about treating the statistical findings of Stud.ᵒˢ II - IV as descriptive and tentative. Clearly, substantial effort will be needed to solve these problems. Although such work has begun (Winne, 1983a), the range of issues concerning ways to analyze data from cognitive mediational research is very large and quite unexplored.

Recapitulation

There is no lack of problems to be conquered in making cognitive mediational research on teaching more valid. The standard trade-offs among internal validity and representativeness, redundancy of data and intrusiveness of data collection, and the like, infect the research characteristic of this paradigm just as much as that characteristic of the process-product paradigm. Nonetheless, the advantage of being able to test cognitive explanations of teaching effects that is offered by a cognitive mediational perspective is substantial. On the strength of this attraction, solving or lessening the problems posed by cognitive mediational research on teaching seems well worth the effort.

As one of the first integrated series of studies reflecting this approach to research on teaching, these problems were difficult to forecast prior to experiencing them. Hence, the substantive weight of our conclusions is not as significant as we had hoped. Exploratory programs of research probably should not be expected to yield substantive breakthroughs. However, the project does have some important findings that can help structure future activities in the field. We turn to those implications in the next section.

IMPLICATIONS

This project offers implications for several audiences. In this section, we discuss some of these implications for three significant groups -- researchers, teacher educators, and teachers.

Implications for Researchers

In the preceding section, problems were identified that are associated with research on teaching carried out within the cognitive mediational paradigm. Obviously, two implications for researchers arising from that section are that those problems need to be articulated more clearly, and that solutions to these better-framed problems should be sought. In Kuhn's (1970) sense, this work will constitute the puzzle-solving activities characteristic of normal scientific development within a paradigm. But the features of this project also have other implications for researchers that are more general than those related to solving the fairly specific set of problems identified previously.

The sort of research illustrated in this project represents two fundamental shifts in the philosophical foundations of research on teaching. In process-product research, relationships between a specific teacher behavior under investigation, such as praise, and student achievement were taken as constant. By this we mean that an a priori, operationally defined instructional stimulus was deemed to have only one effect on students' cognitive mediation of instruction which, in turn, either did or did not have a corresponding effect on how students responded to test items. Brophy's (1981) recent review of process-product research on teacher praise demonstrates the prevalence of this assumption and its untenability (see also Winne, 1982a, for a generalized argument).

Study I of this project demonstrates that a single instructional stimulus can have multiple effects on students' cognitive mediations. We hypothesize also that different instructional stimuli can have the same effect on students' cognitive mediation. Moreover, depending on a particular student's assimilation of information presented over time, whether

the period of time is an episode in a lesson or a semester, we hypothesize that relations between instructional stimuli and cognitive mediations can vary for that one individual. In short, relations between instructional stimuli and students' cognitive mediations are variable over time and context rather than constant. This change in assumptions regarding the relationship between teaching and learning will require recasting models of instruction that grew from earlier research traditions. Specifically, teaching events will need to be viewed as conditions which establish opportunities for students to pursue cognitive operations i.e., as intentional acts (see Scheffler, 1960), rather than as causes of singular cognitive operations that students undertake. The philosophical, methodological, and practical ramifications of this shift need to be developed.

A second basic shift in the philosophy of research on teaching that links to the first is that cognitive activities in which students engage are now the loci for determination of instructional effects. This shift places the agency or mechanism of change completely within a hypothetical construct, i.e., a cognitive operation such as rehearsing, rather than allowing it to straddle the fence of observability in the form of a teacher behavior-plus-cognitive operation that causes achievement. As with the first shift in assumptions, the implications of this relocation of causes of observed instructional effects need to be spelled out as regards such basic issues as the internal validity of experiments and the interpretation of effect size statistics.

Beyond these basic philosophical questions, this project also points to a series of extensions to the methodological suggestions made by Berliner (1976), Dunkin and Biddle (1974), and Snow (1974). We adopt two of Berliner's four categories to describe these extensions. (The methodological and statistical categories were treated earlier, so they are excluded here.)

<u>Dependent variables</u>. The content validity of an achievement test typically has been indicated by the proportion of its items for which information needed to answer an item was presented during lessons. The cognitive mediational view will add to this notion a need to measure the correspondence between cognitive

operations students use to respond to instructional stimuli and cognitive operations students engage to respond to test items. This type of consideration was exemplified in Studies II and III. This additional measure, anchored in behavioral indicator data, also will help to operationalize aspects of transfer that have heretofore been vaguely gauged in research on teaching (cf. Mayer, 1979).

Also, the degree to which an achievement test reflects the effects of instruction is augmented under the cognitive mediational paradigm. Some prior studies of research on teaching have been criticized because they used standardized tests to gauge effects in short-term studies. Researchers have responded by constructing curriculum-specific tests. By adding cognitive mediating strategies to the curriculum, such as hypothesis formation (see Marx & Winne, 1981), concern for the sensitivity of achievement tests now expands from the single dimension concerning content covered to include a second dimension of cognitive mediational strategies (cf. Gagne & Beard, 1978). This is partly illustrated in our research by the measures of training effects employed in Studies II - IV.

Finally, the constant call for multivariate dependent variables can be responded to in part by developing scoring schemes like those we used to score essay test items in Studies II-IV. Extensions of these procedures to other dependent variables defined simultaneously by content, cognitive mediating strategy, and format of the test item (a retrieval task stimulus corresponding in kind to instructional stimuli that students experience during lessons) will need to be developed.

Independent variables. In future research, but especially in non-intervention studies where teachers' delivery of instructional stimuli is uncontrolled by a research design, "independent" variables should be considered from two distinct perspectives. The traditional perspective corresponds to instruction as the researcher sees it, the nominal instructional stimuli of previous process-product research studies. The second perspective applies the criterion of the cognitive function of instructional events to define instructional stimuli. Whereas the former is an a priori

stance for defining hypotheses, the latter is an a posteriori procedure for defining effects. The contrast between these positions, highlighted in Study I, will help to describe the construct validity of putative causes (Cook & Campbell, 1979) or conditions of instructional effectiveness.

A problem with the vast majority of prior experimental research has been its focus on only one or a very small set of instructional events that were defined a priori and only in terms of teacher activities (e.g., Gall, Ward, Berliner, Cohen, Winne, Elashoff, & Stanton, 1978; for contrasts, see Clark et al., 1979 and Good & Grouws, 1979). The possibility rarely was considered that other, undefined instructional stimuli surrounding the treatment variable(s) operationalized in a study also influenced students' cognitive mediations and achievement. For instance, in a review of experiments studying the effects of higher cognitive questions on achievement, this incompleteness of operational definitions for treatment factors led Winne (1979) to ask, "What was the treatment?" The inclusion of behavioral indicator data in studies where students respond to any instructional stimuli they perceive, an extension of the procedure operationalized in our Study IV, will permit treatments to be characterized both functionally and more fully. This also will enlist participants in the instructional acts of lessons, i.e., all the students, to serve as observers of the fidelity of implementation of treatment conditions.

Summary. Addressing the problems uncovered and described by this project should assist in the pursuit of theories of teaching. These theories will be interrelated by their use of students' cognitions as the basis for explaining how teaching and learning interact to affect achievement. By recasting the role attributed to instructional stimuli in the instructional process and by improving methodology for observing students' cognitive mediations, future research also should be able to formulate research questions much more precisely than was possible before.

Implications for Teacher Educators.

The upsurge of cognitive psychological views of teaching and learning has the potential to exert a substantial influence on

both the curriculum of teacher education programs (see Marx & Winne, 1981) and on practices for supervising pre-service teachers. Although it would not be prudent to over-extend the findings of our studies by immediately proposing strong prescriptions for teacher education, there are several suggestions that can be offered.

A major component of teacher education programs focuses on providing pre-service teachers with competence to use an array of discrete teacher skills and to orchestrate these into strategies. This project makes clear the need for pre-service teachers to learn how to examine the ways that these teaching skills and strategies actually influence students since these influences cannot always be predicted reliably. Behavioral indicators can help satisfy this need. As well, student teachers will need preparation in means for analyzing the cognitive operational requirements of tasks that are posed for their pupils by teaching. Conceptually, this parallels the preparation they may receive in task analysis, i.e., analyzing the requirements of tasks by which students demonstrate achievement (Shavelson, 1981). Learning to design behavioral indicators will require pre-service teachers to give thought to how pupils are intended to cognitively mediate teaching skills and strategies.

One further implication for teacher education arises from this project. Supervisors of pre-service teachers should be able to provide much more penetrating analyses of teaching if they, too, make use of data from behavioral indicators of pupils' cognitive mediations. Such data about teaching will help to anchor suggestions for improvement, as well as describe the degree of "instructional rapport" that heretofore has been described less directly by high-inference ratings of pupils' involvement, engaged time, and interest.

Implications for Teachers

Because of the newness of cognitive mediational research on teaching, it is likely that practicing teachers probably are not much more informed about this view than are pre-service teachers other than intuitively by virtue of their experience. Therefore, it seems reasonable to expect that the implications just mentioned

for teacher education generalize to practicing teachers. Two other implications can be mentioned for teachers.

A review of in-service programs by Joyce and Showers (1981) showed that teachers are much more able to practice newly acquired skills and strategies when they receive support for doing so (what these authors labels "coaching"). If, while trying out new teaching methods, ~~teachers can involve their pupils by having them~~ generate behavioral indicators of cognitive mediations, pupils may be able to be the best coaches of all. They will know whether they are confused, or whether a new procedure clicks for them. Hence, by taking both explicit and operational account of students' cognitive mediations of instruction, teachers may be able to obtain the most relevant and timely coaching possible for practicing new teaching methods.

The second implication of this project for teachers concerns remediation and individualization. Explicit consideration of the kinds of cognitive mediations students employ can serve two purposes. One of these is to help the teacher pinpoint instructional stimuli that students are misinterpreting. By supplementing regular teaching with training procedures like those we used in Studies II - IV, these difficulties may be lessened or eliminated. The second purpose that can be served is that of matching instructional delivery to particular cognitive mediational strengths that students may exhibit. Following this suggestion might mean that, at specific points in lessons, a teacher would provide alternative instructional stimuli to different groups of students. Combinations of these two approaches should help the teacher to deal more effectively with the range of individual differences that any group of students presents, and to capitalize on cognitive mediational strengths that might exist in a class.

REFERENCES

Anderson, R.C. Learning in discussions: A resume of the authoritarian-democratic studies. Harvard Educational Review, 1959, 29, 201-215.

Aschner, M.J., Gallagher, J.J., Perry, J., Asfar, S., Jenne, W., & Faar, H. A system for classifying thought processes in the context of classroom verbal interaction. Urbana, IL: University of Illinois, 1965.

Bassett, R.E., & Kibler, R.J. Effect of training in the use of behavioral objectives on student achievement. Journal of Experimental Education, 1975, 44(2), 12-16.

Bellack, A.A., Kliebard, H.M., Hyman, R.T., & Smith, F.L. The language of the classroom. New York: Teachers College Press, 1966.

Berliner, D.C. Impediments to the study of teacher effectiveness. Journal of Teacher Education, 1976, 27, 5-13.

Bloom, B.S., Engelhart, M.B., Furst, E.J., Hill, W.H., and Krathwohl, D.R. Taxonomy of educational objectives. The classification of educational goals. Handbook I: Cognitive domain. New York: Longmans Green, 1956.

Bronfenbrenner, U. The experimental ecology of education. Educational Researcher, 1976, 5(9), 5-15.

Brophy, J.E. Teacher praise: A functional analysis. Review of Educational Research, 1981, 51, 5-32.

Brown, A.L. Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), Advances in instructional psychology, vol. 1. Hillsdale, NJ: Lawrence Erlbaum, 1978.

Brown, J.S., & Burton, R.R. Diagnostic models for procedural bugs in basic mathematical models. Cognitive Science, 1978, 2, 155-192.

Calfee, R. Cognitive psychology and educational practice. In D.C. Berliner (Ed.), Review of research in education, vol. 9. Washington, D.C.: American Educational Research Association, 1981.

Clark, C.M., Gage, N.L., Marx, R.W., Peterson, P.L., Stayrook, N.G., & Winne, P.H. A factorial experiment on teacher structuring, soliciting, and reacting. Journal of Educational Psychology, 1979, 71, 534-552.

Cohen, J., & Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Erlbaum, 1975.

Cook, T.D., & Campbell, D.T. Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally, 1979.

Corno, L. Cognitive organizing in classrooms. Curriculum Inquiry, 1981, 11, 359-377.

Dansereau, D.F., Collins, K.W., McDonald, B.A., Holley, C.D., Garland, J., Diekhoff, G., and Evans, S.H. Development and evaluation of a learning strategy training program. Journal of Educational Psychology, 1979, 79, 64-73.

Dillon, J.T. Cognitive correspondence between question/statement and response. American Educational Research Journal, 1982, 19, 540-551.

Doyle, W. Student mediating responses in teaching effectiveness (Final report). Denton, TX: North Texas State University, 1980.

Doyle, W. Paradigms for research on teaching. In L.S. Shulman (Ed.), Review of research in education, vol. 5. Itasca, IL: Peacock, 1978.

Dunkin, M.J., & Biddle, B.J. The study of teaching. Holt, Rinehart, & Winston, 1974.

Ericsson, K.A., & Simon, H.A. Verbal reports as data. Psychological Review, 1980, 87, 215-251.

Flanders, N.A. Helping teachers change their behavior (Final report). Ann Arbor, MI: University of Michigan, School of Education, 1963.

Flanders, N.A. Teacher influence, pupil attitudes, and achievement. Minneapolis, MN: University of Minnesota (U.S. Office of Education Cooperative Research Project No. 397), 1960. (Mimeographed)

Gage, N.L. The scientific basis of the art of teaching. New York: Teachers College Press, 1978.

Gage, N.L. Theories of teaching. In E.R. Hilgard (Ed.), Theories of learning and instruction. Chicago: University of Chicago Press, 1964.

Gage, N.L. Paradigms for research on teaching. In N.L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.

Gage, N.L., & Winne, P.H. Performance-based teacher education. In K. Ryan (Ed.), Teacher education. Chicago: University of Chicago Press, 1975.

Gage, N.L. The causal connections underlying teacher education: Their status and prospects. Paper presented at the annual meeting of the American Educational Research Association, Boston, 1980.

Gagné, E.D. Long-term retention of information following learning from prose. Review of Educational Research, 1978, 48, 629-665.

Gagné, R.M., & Beard, J.G. Assessment of learning outcomes. In R. Glaser (Ed.), Advances in instructional psychology, vol. 1. Hillsdale, NJ: Lawrence Erlbaum Associates, 1978.

Gagné, R.M. The conditions of learning, 3 ed. New York: Holt, Rinehart, and Winston, 1977.

Gall, M.D., & Gall, J.P. The discussion method. In N.L. Gage (Ed.), The psychology of teaching methods. Chicago: University of Chicago Press, 1976.

Gall, M.D., Ward, B.A., Berliner, D.C., Cahen, L.S., Winne, P.H., Elashoff, J.D., & Stanton, G.S. Effects of questioning techniques and recitation on student learning. American Educational Research Journal, 1978, 15, 175-199.

Good, T.L., & Grouws, D. The Missouri Mathematics Effectiveness Project: An experimental study in fourth-grade classrooms. Journal of Educational Psychology, 1979, 71, 355-362.

Hilgard, E.R. Theories of learning, 2nd ed. New York: Appleton-Century-Crofts, 1956.

Hunkins, F.P. Questioning strategies and techniques. Boston: Allyn and Bacon, 1972.

Joyce, B.R., & Showers, B. Transfer of training: The contribution of coaching. Journal of Education, 1981, 163, 163-72.

Kaplan, A. The conduct of inquiry: Methodology for behavioral science. San Francisco: Chandler, 1964.

Koopman, C., & Newston, D. Level of analysis in the perception of ongoing instruction. Journal of Educational Psychology, 1981, 73, 212-223.

Kuhn, T.S. The structure of scientific resolutions, 2nd Ed. Chicago: University of Chicago Press, 1970.

Marx, R.W. Teacher judgment of students' cognitive and affective outcomes (Doctoral dissertation, Stanford University, 1978). Dissertation Abstracts International, 1978, 39(6), 3475-A. (Order No. 7822540)

Marx, R.W., & Winne, P.H. The shape of teaching to come: Implications for the National Teacher Examination. Princeton, NJ: National Teacher Examination Policy Council, Educational Testing Service, 1981.

Marx, R.W., Winne, P.H., and Howard, D.C. Influence of cognitive style on students' perception of instruction (Final report). Burnaby, B.C.: Instructional Psychology Research Group, Simon Fraser University, 1982.

Mayer, R.E. Can advance organizers influence meaningful learning? Review of Educational Research, 1979, 49, 371-383.

Morine-Dershimer, G. Pupil perceptions of teacher praise. Elementary School Journal, 1982, 82, 421-434.

Morine-Dershimer, G., & Fagal, F. Why do you ask? Paper presented at the annual meeting of the American Educational Research Association, Boston, 1980.

Nisbett, R.E., & Wilson, T.D. Telling more than we can know: Verbal reports on mental processes. Psychological Review, 1977, 84, 231-259.

Peterson, P.L. Interactive effects of student anxiety, achievement orientation, and teacher behavior on student achievement and attitude. Journal of Educational Psychology, 1977, 69, 779-792.

Peterson, P.L., Swing, S.R., Braverman, M.T., & Buss, R. Students' aptitudes and their reports of cognitive processing during instruction. Journal of Educational Psychology, 1982, 74, 535-547.

Peterson, P.L., & Walberg, H.J. (Eds.). Research on teaching. Berkeley, CA: McCutchan, 1979.

Program on Teaching Effectiveness. A factorially designed experiment on teacher structuring, soliciting and reacting. (Research and Development Memorandum No. 147). Stanford, CA: Stanford Center for Research and Development in Teaching, 1976.

Rice, J. The futility of the spelling grind. Forum, 1897, 23, 163-172, 409-419.

Rosenshine, B., & Furst, N. The use of direct observation to study teaching. In R.M.W. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand McNally, 1973.

Rothkopf, E.Z. Writing to teach and reading to learn: A perspective on the psychology fo written instruction. In N.L. Gage (Ed.), The psychology of teaching methods. Chicago: University of Chicago Press, 1976.

Ryan, T. W., Hildyard, A., Jantzi, D., and Pike, R. The IEA classroom environment study (Interim Report). Toronto: Ontario Institute for Studies in Education, January, 1983.

Scheffler, I. The language of education. Springfield, IL: Thomas, 1960.

Shavelson, R.J. Teaching mathematics: Contributions of cognitive research. Educational Psychologist, 1981, 16, 23-44.

Shavelson, R.J., & Stern, P. Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. Review of Educational Research, 1981, 51, 455-498.

Simon, A., & Boyer, E.G. (Eds.). Mirrors for behavior: An anthology of classroom observation instruments. Supplementary vols. A and B. Philadelphia: Research for Better Schools, 1970.

Skinner, B.F. Science and human behavior. New York: MacMillan, 1953.

Smith, H.A., Beugger, P., Frazier, R., Magnoli, M., & Wied, A. Exploring Living Things: Laidlaw Exploring Science Program. Toronto: Doubleday, 1977.

Snow, R.E. Designs for research on teaching. Review of Educational Research, 1974, 44, 265-291.

Snow, R.E. Theory construction for research on teaching. In R.M.W. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand McNally, 1973.

Snow, R.E. Brunswikian approaches to research on teaching. American Educational Research Journal, 1968, 5, 455-489.

Stayrook, N.G., Corno, L., and Winne, P.H. Path analyses relating student perceptions of teacher behavior to student achievement. Journal of Teacher Education, 1973, 29(2), 51-56.

Stecher, A., et al. Your senses. Toronto: Holt, Rinehart, and Winston of Canada, 1976.

Thorndike, R.L., & Hagen, E. Examiner's manual: Cognitive Abilities Test. Boston: Houghton-Mifflin, 1971.

Tobias, S. When do instructional methods make a difference? Educational Researcher, 1982, 11(4), 4-9.

Weinstein, R. Student perceptions of schooling. Elementary School Journal, in press.

Weinstein, R.S., Middlestadt, S.E., Brattesani, K.A., & Marshall, H.H. Student perceptions of differential treatment. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Winne, P.H. Methodology for theoretical research about teaching, student cognition and achievement. Unpublished draft, 1983. (a)

Winne, P.H. Cognitive processing in the classroom. In T. Husen and T.N. Postlethwaite (Eds.), International encyclopedia of education. Oxford: Pergamon Press, 1983, in press. (b)

Winne, P.H. Distortions of construct validity in multiple regression analyses. Manuscript submitted for publication, 1983. (c)

Winne, P.H. Minimizing the black box problem to enhance the validity of theories about instructional effects. Instructional Science, 1982, 11, 13-28. (a)

Winne, P.H. Training students to process text with adjunct aids. Manuscript submitted for publication, 1982. (b)

Winne, P.H. Experiments relating teachers' use of higher cognitive questions to student achievement. Review of Educational Research, 1979, 44, 13-49.

Winne, P.H., & Marx, R.W. Students' and teachers' views of thinking processes for classroom learning. Elementary School Journal, 1982, 82, 493-518.

Winne, P.H., & Marx, R.W. Matching students' cognitive responses to teaching skills. Journal of Educational Psychology, 1980, 72, 257-264.

Winne, P.H., & Marx, R.W. Perceptual problem solving. In P.L. Peterson and H.J. Walberg (Eds.), Research on teaching. Berkeley, CA: McCutchan, 1979

Winne, P.H., & Marx, R.W. Reconceptualizing research on teaching. Journal of Educational Psychology, 1977, 69, 668-678.